Express Mail Label No. EV780362226US
Date of Deposit: November 16, 2005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Keith D. Allen	Examiner:	Qian, Celine X
Serial No.:	10/005,467	Group Art Unit:	1636
Filed:	December 4, 2001	Docket No.:	R758/75658.295
Confirmation No.	7217		
Title:	Transgenic Mice Containing PTP36 Tyrosine Phosphatase Gene Disruptions		

Mail Stop Appeal Brief
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF

Sir:

In regard to the above-referenced application, Appellant submits this Appeal Brief.

I. REAL PARTY IN INTEREST

Deltagen, Inc., the assignee, is the real party in interest. The right of Deltagen, Inc. to take action in the subject application was established by virtue of the assignment from the inventor, Keith D. Allen, to Deltagen, Inc., recorded at Reel 012697, Frame 0388.

II. RELATED APPEALS AND INTERFERENCES

As of the filing date of this appeal, the undersigned legal representative of the Appellant confirms there are no prior or pending appeals, interferences or

judicial proceedings known to Appellant, the Appellant's legal representative, or assignee which are related to, directly affect or are directly affected by or have a bearing on the Board's decision in the present appeal.

III. STATUS OF CLAIMS

Claims 1-27, 33-36, 38-46, 48-52 and 58 have been canceled, and claims 28-32, 37, 47 and 53-57 are pending. Claims 28-32, 37, 47 and 53-57 are appealed.

IV. STATUS OF AMENDMENT

An after final Supplementary Amendment was filed September 8, 2005 amending claims 28-32 and 55-57, and canceling claims 52 and 58. This amendment was entered by the Examiner. Following entry of this amendment, claims 28-32, 37, 47 and 53-57 are pending and rejected.

V. SUMMARY OF CLAIMED SUBJECT MATTER

Claim 28, the sole independent claim, is drawn to a transgenic mouse whose genome comprises a null allele in the endogenous PTP36 gene. A ***transgenic mouse*** is a mouse whose genome has been modified such that a specific gene (e.g., the PTP36 gene) has been disrupted by the method of gene targeting. Specification, page 6, lines 29-30. The term "***gene targeting***" refers to a type of homologous recombination that occurs when a fragment of exogenous DNA (e.g., a targeting vector; see Figure 4) is introduced into a mammalian cell and that exogenous DNA locates and recombines with the endogenous homologous sequences. Specification, page 5, lines 23-25. A ***null allele*** is

disruption of the gene in which there is no significant expression of the PTP36 gene. Specification, page 6, lines 23-24. A normal or wild-type mouse has two (2) copies of the PTP36 gene, each referred to as a ***wild-type allele***. The claimed transgenic mouse includes both the heterozygous animal (i.e., one defective PTP36 allele and one wild-type PTP36 allele) and the homozygous animal (i.e., two defective PTP36 alleles). *Id.*, page 6, lines 28- page 7, lines 2. By disrupting or "***knocking out***" the gene, the function of the PTP36 gene can be determined by comparing the transgenic mouse with a wild-type control mouse. *Id.*, page 14, lines 5-10.

Claim 37 is drawn to a cell or tissue isolated from the transgenic mouse of claim 28. Specification, page 2, lines 20-23; page 32, lines 22-27.

Claim 47 is drawn to a method of producing the transgenic mouse of claim 28. The method comprises the steps of introducing a targeting construct capable of disrupting an endogenous PTP36 allele into a mouse embryonic stem cell; selected for the mouse embryonic stem that has undergone homologous recombination; introducing the mouse embryonic stem cell selected for into a blastocyst; implanting the resulting blastocyst into a pseudopregnant mouse, wherein the resultant mouse gives birth to a chimeric mouse; and breeding the chimeric mouse to produce the transgenic mouse. Specification, page 8, line 10- page 14, line 4.

Claim 53, dependent from claim 28, is drawn to the transgenic mouse of claim 28 wherein the mouse is heterozygous for the null allele. Specification, page 14, lines 1-2; page 48, line 28- page 49, line 2; page 50, lines 10-11.

Claim 54, dependent from claim 28, is drawn to the transgenic mouse of claim 28 wherein the mouse is homozygous for the null allele. Specification,

page 3, lines 5-6; specification, page 6, line 28- page 7, line 2; page 12, lines 1-4; page 49, lines 1-27. The inherent properties of the homozygous mouse of claim 54 are recited in dependent claims 29-32.

Claim 55, dependent from claim 28, is drawn to the transgenic mouse of claim 28 wherein the null allele comprises a gene encoding a selection marker. A *"selection marker"* refers to a gene encoding a product that enables only the cells that carry the gene to survive and/or grow under certain conditions. Specification page 5, lines 3-8.

Claim 56 is drawn to the transgenic mouse of claim 55 wherein the gene encoding a selection marker is a neomycin resistance gene. The neomycin gene confers resistance to the compound G418 such that cells that do not carry the neomycin resistance gene are killed by G418. Specification, page 7, lines 3-9; Figure 4.

Claim 57 is drawn to the transgenic mouse of claim 56 wherein the null allele further comprises a LacZ gene. See Figure 4; specification page 50, line 10- page 54, line 9. The LacZ gene allows one to determine where the PTP36 gene is expressed in heterozygous mice by staining with the X-Gal substrate. Specification, page 50, lines 10-13.

Claims 29-31 are each drawn to the transgenic mouse of claim 54 wherein the mouse exhibits a specific uterine abnormality, relative to a wild-type control mouse. Specification page 49, lines 14-23. The uterine abnormality phenotype is an inherent property of the homozygous mouse of claim 54.

Claim 32 is drawn to the transgenic mouse of claim 54 wherein the mouse exhibits, relative to a wild-type control mouse, increased organ weight in one or more of the specified organ systems. Increased organ weight in one or more of

the specified organ systems is an inherent property of the homozygous mouse of claim 54.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

A. Whether Claims 28-32, 37, 47 and 53-57 are Unpatentable Under 35 U.S.C. § 101

Claims 28-32, 37, 47 and 53-57 are rejected under 35 U.S.C. § 101 on the grounds that the claimed invention is allegedly not supported by either a specific or substantial asserted utility or a well-established utility.

B. Whether Claims 28-32, 37, 47 and 53-57 are Unpatentable Under 35 U.S.C. § 112, first paragraph: Enablement

Claims 28-32, 37, 47 and 53-57 are rejected under 35 U.S.C. § 112, first paragraph, on the grounds that the patent application does not enable one skilled in the art to use the claimed invention as a result of the alleged lack of either a specific or substantial asserted utility or a well-established utility.

C. Whether Claims 28-32, 37, 47, 53-54, and 57 are Unpatentable Under 35 U.S.C. § 112, first paragraph: Written Description

Claims 28-32, 37, 47, 53-54, and 57 are rejected under 35 U.S.C. § 112, first paragraph, on the grounds that the claims contain subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor, at the time of filing, had possession of the claimed invention.

D. Whether Claim 32 is Unpatentable Under 35 U.S.C. § 112, second paragraph: Definiteness

Claim 32 is rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to point out and distinctly claim the subject matter which applicant regards as the invention.

VII. ARGUMENT

A. Technology Background

The following steps are performed to obtain transgenic mice within the scope of the claims *i.e.*, transgenic mice whose genomes comprise a null allele in the endogenous PTP36 gene. It is assumed that the Board is familiar with the basic principles of molecular genetics.

1. A targeting vector is prepared. The targeting vector is a DNA construct that includes a LacZ-Neo cassette flanked on either side by sequences that are identical to sequences in the endogenous mouse PTP36 gene. These flanking sequences are referred to as "***homology arms***" and are depicted in Figure 4 of the application as including SEQ ID NO: 2 and SEQ ID NO: 3. The LacZ-Neo cassette includes a LacZ gene (which encodes the bacterial enzyme β-galactosidase) and a Neo gene (which encodes a protein that confers resistance to the drug neomycin). The LacZ gene in the cassette lacks a promoter (the sequences required to direct transcription of the LacZ gene), whereas the Neo gene in the cassette is operatively linked to a promoter. See Figure 4.

2. The targeting vector is introduced into mouse ***embryonic stem (ES)*** cells. ES cells are pluripotential cells isolated from mouse embryos that may be grown and manipulated in culture. The homology arms of the targeting vector

pair up with the complementary sequence in the endogenous mouse PTP36 gene. Through a process known as "***homologous recombination,***" the targeting vector recombines with the PTP36 gene, precisely replacing specific bases of one chromosomal copy of the PTP36 gene with the LacZ-Neo cassette (the replaced bases of the PTP36 gene are depicted in Figure 2A (underlined nucleotides) and also in Figure 3 of the application). The replacement of these bases with the LacZ-Neo cassette ablates the function of the gene. The resulting disrupted PTP36 gene is referred to as a "***null allele.***" ES cells are then screened for resistance to the drug neomycin. If an ES cell is neomycin-resistant, this indicates that the ES cell expresses the Neo gene and, therefore, the ES cell has the LacZ-Neo cassette inserted into its genome. Disruption of the PTP36 gene in neomycin-resistant ES cells is then confirmed using the Polymerase Chain Reaction (PCR) technique and Southern blot analysis.

3. ES cells in which the PTP36 gene has been disrupted by homologous recombination are then introduced into a host mouse blastocyst (a preimplantation embryo consisting of a hollow sphere of cells) obtained from the C57BL/6 mouse strain. C57BL/6 mice carry the black coat color allele. The ES cells are derived from the 129/OlaHsd mouse strain. 129/OlaHsd mice carry the agouti coat color allele. The manipulated blastocyst is then implanted into a pseudopregnant mouse (a female mouse that has been mated with a vasectomized male); the blastocyst develops and the female mouse gives birth to a ***chimeric mouse*** that developed from the manipulated blastocyst. Chimeric mice have regions of agouti colored fur—indicating tissue derived from the implanted ES cells—and regions of black fur—indicating tissue derived from the C57BL/6 host blastocyst.

4. Chimeric male mice are then mated with C57BL/6 females. If a chimeric male mouse has a germ line that is derived from an ES cell, then sperm from the chimeric mouse will carry the agouti coat allele found in the 129/OlaHsd strain. The agouti coat allele is dominant over the black coat allele. As a result, if a chimeric male mouse has a germ line derived from an ES cell, then the progeny of that mouse will all have pure agouti coats. The agouti progeny, termed the F1 generation, receive one wild-type copy of the PTP36 gene from the C57BL/6 mother, and one copy of the PTP36 gene from the ES cell-derived germ line of their chimeric fathers. Because only one copy of the PTP36 gene is knocked out in the ES cells, half of the F1 progeny inherit a disrupted copy of the PTP36 gene—a null allele—from their father, and half inherit a normal (wild-type) copy. In genetic nomenclature, the wild-type copy of the PTP36 gene is referred to as "+" and the null allele of the PTP36 gene is referred to as "-." Hence, the F1 progeny are either +/+ or +/-. Mice that are +/- are referred to as "***heterozygous mice***" or as "***heterozygotes***" and are identified by genotyping techniques (Southern blot analysis and PCR analysis).

5. F1 heterozygous mice are then backcrossed to C57BL/6 mice, and heterozygous male and female mice in the resulting litters are again selected. These heterozygous males and females are then intercrossed. The resulting mice are referred to as the F2 generation. One quarter of the F2 progeny are +/+, one quarter are heterozygotes (+/-) and one quarter have two copies of the null allele (-/-) of the PTP36 gene. Mice that are -/- are referred to as "***homozygous mice***" or as "***homozygotes***." The F2 heterozygous mice and homozygous mice are then phenotypically characterized and compared with age-matched and gender-matched F2 +/+ mice, referred to herein as "***wild-type control mice***." It typically

takes approximately one year from the initial targeting of the gene in ES cells to generate F2 homozygous mice.

6. The expression pattern of the LacZ gene is determined in tissue samples from heterozygous mice. Because the LacZ gene does not have its own promoter, transcription of the LacZ gene is driven by the endogenous PTP36 gene promoter. As a result, the LacZ gene is expressed in the same cells that the PTP36 gene is expressed in. LacZ expression is detected by contacting tissue from heterozygous mice with a chemical named "X-gal." β-galactosidase (the enzyme encoded for by the LacZ gene) catalyzes the conversion of X-gal into a blue colored reaction product; the staining of tissue with this blue product is visualized using a microscope.

B. Legal Standard

1. General Principles

(a) Utility under 35 U.S.C. § 101

35 U.S.C. § 101 states that "[w]hoever invents... any new and ***useful*** ...composition of matter ...may obtain a patent therefore...." 35 U.S.C. § 101 (emphais added). In *Brenner v. Manson,* 383 U.S. 519 (1966), the Supreme Court explained that

> [t]he basic quid pro quo contemplated by the Constitution and the Congress for granting a patent monopoly is the benefit derived by the public from an invention with ***substantial*** utility. Unless and until a process is refined and developed to this point – where ***specific*** benefit exists in currently available form – there is insufficient justification for permitting an applicant to engross what may prove to be a broad field.

Brenner, 383 U.S. at 534-535 (emphases added). This language has subsequently been interpreted by the Federal Circuit to mean that an invention is

useful if it has both a "***substantial***" utility and a "***specific***" utility. *See In re Fisher*, 421 F.3d 1365, 1371 (Fed. Cir. 2005).

The terms "***practical utility***" and "***real world utility***" are often used in determining whether an invention offers a "substantial" utility. *See Nelson v. Bowler*, 626 F.2d 853, 856, 206 USPQ 881, 883 (C.C.P.A. 1980). An invention has a practical and real world utility, and hence a substantial utility, when "one skilled in the art can use a claimed discovery in a manner which provides some immediate benefit to the public." *In re Fisher*, 421 F.3d at 1371 (citing *Nelson v. Bowler*, 626 F.2d 853, 856). Accordingly, for an invention to have a substantial utility, "an asserted use must show that the claimed invention has a significant and presently available benefit to the public." *In re Fisher*, 421 F.3d at 1371.

With regard to the "specific" utility requirement, "an application must disclose a use which is not so vague as to be meaningless." *Id.* An asserted use must show that the "claimed invention can be used to provide a well-defined and particular benefit to the public." *Id.* A specific utility is particular to the subject matter claimed as compared to a "general utility" that would be applicable to the broad class of the invention. *See id; see also* U.S. Pat. & Trademark Off., Manual of Patent Examining Procedure § 2107.01 (8th ed. 2001, rev. May 2004).

The Federal Circuit has recognized that "[t]he threshold of utility is not high." *Juicy Whip v Orange Bang*, 185 F.3d 1364, 1366 (Fed. Cir. 1999). Rather, an invention has utility "if it is capable of providing some identifiable benefit." *Id.* A finding of lack of utility is warranted only if the invention is "totally incapable of achieving a useful result." *Brooktree Corp. v. Advanced Micro Devices*, Inc., 977 F.2d 1555, 1571 (Fed. Cir. 1992). Accordingly, utility must be found even if an invention is only ***partially*** successful in achieving a ***single*** credibly asserted use.

Moreover, utility exists even if further research and development on the claimed invention is expected. *See In re Brana*, 51 F.3d 1560, 1568 (Fed. Cir. 1995).

An assertion of utility is presumed to be true. *See In re Brana*, 51 F.3d at 1566 (citing *In re Marzocchi*, 439 F.2d 220, 223 (C.C.P.A. 1971)). From this it follows that the PTO has the initial burden of challenging as non-credible a presumptively correct assertion of utility in the disclosure. Only after the PTO provides evidence showing that one of ordinary skill in the art would reasonably doubt the credibility of the asserted utility does the burden shift to the applicant to provide rebuttal evidence sufficient to convince such a person of the invention's asserted utility. *See id.* If one skilled in the art would "immediately appreciate why" a claimed invention has a specific and substantial utility, then the invention has a "well-established utility," notwithstanding the failure of the application to explicitly state that utility. *See* U.S. Pat. & Trademark Off., Manual of Patent Examining Procedure § 2107.02 II.B.

Any reasonable use that an applicant has identified for the invention that can be viewed as providing a public benefit should be accepted as a substantial utility. *See* U.S. Pat. & Trademark Off., Manual of Patent Examining Procedure § 2107.01. Proof of an asserted utility "beyond a reasonable doubt" or as a statistical certainty is not required; instead, evidence will be sufficient if, considered as a whole, it leads one skilled in the art to conclude that the asserted utility is more likely than not true. *See* U.S. Pat. & Trademark Off., Manual of Patent Examining Procedure § 2107.02 VII. The character and amount of evidence needed may vary depending on what is claimed. *See In re Jolles*, 628 F.2d 1322, 1325 (C.C.P.A. 1980) (citing *In re Chilowsky* 229 F.2d 457, 461 (C.C.P.A. 1956)). Nonetheless, there is "no basis in the statutes or decisions for requiring

any more conclusive evidence of operativeness in one type of case than another." *In re Chilowsky*, 229 F.2d at 461-2.

Proof of utility is conclusively established as a matter of law if an infringing party has made, sold, or used the claimed invention. *See Raytheon Co. v. Roper*, 724 F.2d 951, 959 (Fed. Cir. 1983). The Federal Circuit has characterized this as a "common sense" rule because "[p]eople rarely, if ever, appropriate useless inventions." *Raytheon Co.* 724 F. 2d at 959. Similarly, proof of utility is supported when a claimed invention has been met with commercial success. *See Raytheon Co. v. Roper Corp.* 724 F. 2d at 959; *see also In re Fisher*, 421 F.3d at 22. This rule is a logical corollary of the Supreme Court's instruction in *Brenner* that the patent system is "related to the world of commerce rather than to the realm of philosophy." *Brenner*, 383 U.S. at 536 (quoting *In re Ruschig*, 343 F.2d 965, 970 (C.C.P.A. 1965)). In other words, just as "[p]eople rarely, if ever, ***appropriate*** useless inventions," *Raytheon Co.* 724 F. 2d at 959 (emphasis added), people rarely, if ever, ***purchase*** useless inventions.

(b) Enablement Under 35 U.S.C. § 112, first paragraph

If an invention lacks utility under 35 U.S.C. § 101, then the patent application also fails to enable one of ordinary skill in the art to use the invention as required by 35 U.S.C. § 112, first paragraph. *See In re Kirk*, 376 F.2d 936, 942 (C.C.P.A. 1976) ("Necessarily, compliance with § 112 requires a description of how to use presently useful inventions, otherwise an applicant would anomalously be required to teach how to use a useless invention."); *see also* U.S. Pat. & Trademark Off., Manual of Patent Examining Procedure § 2107.01.

2. In re Fisher

The Federal Circuit's recent utility analysis in *In re Fisher*, 421 F.3d 1365 (Fed. Cir. 2005) is particularly relevant to the instant patent application. In *In re Fisher*, the Appellant filed a patent application which disclosed more than 30,000 short nucleotide sequences, termed "expressed sequence tags" (ESTs) from *Zea mays* (maize). *Ex parte Fisher*, App. No. 2002-2046, slip op. at 1 (Bd. Pat. App. Int. Mar. 16, 2004). ESTs correspond to fragments of expressed genes. ESTs are routinely and rapidly generated by machine-automated sequencing of cDNA clones at a rate of many hundreds of ESTs per day. At the time of writing, there are almost six hundred thousand publicly-described maize ESTs, and almost six million publicly-described human ESTs. *See* National Center for Biotechnology Information (NCBI) dbEST Database (<<http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html>>).

Upon final rejection of claims drawn to five elected ESTs for lack of a specific and substantial utility and for lack of enablement, Fisher appealed to the Board arguing that ESTs have a number of specific uses, including 1) use for the identification of polymorphisms; and 2) use as probes or primers for the identification of orthologous genes from other plants and organisms. *Ex parte Fisher*, App. No. 2002-2046, slip op. at 1.

The Board affirmed the Examiner's rejection, finding that the asserted uses were insubstantial because the function of the genes from which the ESTs were derived was unknown. *Ex parte Fisher*, App. No. 2002-2046, slip op. at 15. Specifically, the Board found that the asserted use of identifying polymorphisms was insubstantial because without knowing the function of the gene, detection of a polymorphism provides only "the barest information." *Id.* Similarly, using the

ESTs from genes having no known function to isolate other genes having no known function was insubstantial. *See id.*

The Federal Circuit affirmed. The court held first that to constitute a substantial utility, an asserted use "must show that the claimed invention has a significant and presently available benefit to the public." *In re Fisher*, 421 F.3d at 1371. In this case, the Federal Circuit held that "our precedent dictates that the '643 application does not meet the utility requirement of § 101 because Fisher ***does not identify the function for the underlying protein-encoding genes.***" *In re Fisher*, 421 F.3d at 1376 (emphasis added). The court emphasized that without the identification of the function of their cognate genes, the claimed invention has "not been researched and understood to the point of providing an immediate, well-defined, real-world benefit to the public meriting the grant of a patent." *Id.* Instead, the claimed ESTs were merely "objects upon which scientific research could be performed with no assurance that anything useful will be discovered in the end." *In re Fisher*, 421 F.3d at 1373 (quoting *Brenner*, 383 U.S. at 535).

The court focused on the fact that the Appellant provided ***no*** evidence whatsoever—such as test data or declarations—to prove that any of the claimed ESTs had successfully been put to the alleged uses in the real world. This evidentiary failure left the court with "no choice but to conclude that the claimed ESTs do not have a 'substantial' utility under § 101." *In re Fisher*, 421 F.3d at 1374.

The court also held that to constitute a specific utility, an asserted use "must also show that the claimed invention can be used to provide a well-defined and particular benefit to the public." *In re Fisher*, 421 F.3d at 1371. In this

case, the asserted uses were not specific, the court found, because any EST could be used for the asserted uses. *See In re Fisher*, 421 F.3d at 1374.

The court agreed that commercial success may support the utility of an invention. *See In re Fisher*, 421 F.3d at 1377-78 (citing *Raytheon Co. v. Roper Corp.*, 724 F.2d 951, 959 (Fed. Cir. 1983)). In this case, however, there was no evidence that any company had even expressed interest in, let alone purchased, a single one of the claimed ESTs. *See In re Fisher*, 421 F.3d at 1378.

Finally, the court addressed concerns expressed by the government that allowing EST patents without proof of utility would discourage and delay research. The court warned that such public policy concerns must ***not*** be considered by courts in determining whether the utility requirement is met; such concerns should be addressed only by Congress. *See In re Fisher*, 421 F.3d at 1378.

C. The Subject Matter Of Each of Claims 28-32, 37, 47, 53-57 Has Multiple Credible and Well-Established Uses

The following arguments show that the presently-claimed invention has ***multiple*** credible and well-established utilities, each of which is substantial and specific. As a claimed invention need only be ***partially*** successful in achieving a ***single*** credibly asserted or well-established utility in order to pass muster under 35 U.S.C. § 101, *see* Brooktree *Corp.*, 97 F.2d at 1571, the claimed invention more than adequately satisfies the utility requirement.

Note that in the discussion that follows, certain uses are described with reference to the homozygous mouse (a mouse that has two null alleles of the PTP36 gene) and tissue and cells isolated from the homozygous mouse. Homozygous mice, also referred to herein as "PTP36 knockout mice," are within the scope of claims 28-32 and 54-57; isolated tissues and cells from homozygous

mice are within the scope of claim 37; and the inherent phenotypes exhibited by homozygous mice are recited in claims 29-32. Since the homozygous mouse has utility, the heterozygous mouse (a mouse having only one null allele of the PTP36 gene) necessarily has utility because heterozygous mice must be used to generate the homozygous mice. *See* Section VII.A *supra*. Heterozygous mice are within the scope of claims 28, 53, and 55-57. Heterozygous mice, and tissues and cells isolated from heterozygous mice (within the scope of claim 37), are also used to determine the gene expression pattern of PTP36. *See* Section VII.C.5 below. Similarly, because the homozygous mice and the heterozygous mice have utility, the claimed method for generating homozygous and heterozygous mice (claim 47) also has utility.

1. The PTP36 Knockout Mouse Is Used to Study The Function Of The PTP36 Gene.

(a) The Asserted Utility Of Using The PTP36 Knockout Mouse To Study The Function Of The PTP36 Gene Is A Well-Established Utility.

Contary to the Examiner's assertions, the present invention has a well-established utility since a person of ordinary skill in the art "would immediately appreciate why" (M.P.E.P. § 2107.02 II.B) the claimed PTP36 knockout mice are useful. Specifically, knockout mice have the inherent and well-established utility of defining the function and role of the disrupted target gene.

The sequencing of the human genome has produced thousands of genes whose function has yet to be determined. Knockout mice are the recognized "gold standard" for determining the function of those genes. According to the United States National Institutes of Health (NIH), knockout mice are critical to studying gene function:

> Over the past century, the mouse has developed into ***the premier mammalian model system*** for genetic research. Scientists from a wide range of biomedical fields have gravitated to the mouse because of its close genetic and physiological similarities to humans, as well as the ease with which its genome can be manipulated and analyzed.
>
> . . .
>
> In recent decades, researchers have utilized an array of innovative genetic technologies to produce custom-made mouse models for a wide array of specific diseases, ***as well as to study the function of targeted genes***....[P]owerful approaches, dependent on homologous recombination, have permitted the development of tools to ***"knock out" genes***, which involves replacing existing genes with altered versions... To preserve these extremely valuable strains of mice and to assist in the propagation of strains with poor reproduction, researchers have taken advantage of state-of-the-art reproductive technologies, including cryopreservation of embryos, in vitro fertilization and ovary transplantation.

(<<http://www.genome.gov/pfv.cfm?pageid=10005834>>)(emphasis added)(Exhibit A)(see Amendment filed February 17, 2005). Thus, the knockout mouse has been accepted by the NIH as ***the*** premier model for determining and studying gene function.

Knockout mice are so well accepted as useful for determining gene function that the director of the NIH Chemical Genomics Center of the National Human Genome Research Institute (among others, including Capecchi, Bradley, Joyner, Nagy and Skarnes, each a pioneer and recognized expert in the field of mammalian genomics) has proposed creating knockout mice for ***all*** mouse genes:

> Now that the human and mouse genome sequences are known, ***attention has turned to elucidating gene function*** and identifying gene products that might have therapeutic value. The laboratory mouse (*Mus musculus*) has had a prominent role in the study of human disease mechanisms throughout the rich, 100-year history of classical mouse genetics, exemplified by the lessons learned from naturally occurring mutants such as agouti, reeler and obese.... Among the model

> organisms, the mouse offers particular advantages for the study of human biology and disease: (i) the mouse is a mammal, and its development, body plan, physiology, behavior and diseases have much in common with those of humans; (ii) almost all (99%) mouse genes have homologs in humans; and (iii) the mouse genome supports targeted mutagenesis in specific genes by homologous recombination in embryonic stem (ES) cells, allowing genes to be altered efficiently and precisely.
>
> . . .
>
> A coordinated project to systematically knock out all mouse genes is likely to be of enormous benefit to the research community, given ***the demonstrated power of knockout mice to elucidate gene function,*** the frequency of unpredicted phenotypes in knockout mice, the potential economies of scale in an organized and carefully planned project, and the high cost and lack of availability of knockout mice being made in current efforts.

(Austin et al., Nature Genetics (2004) 36(9):921-24, 921)(emphasis added)(Exhibit B) (see Amendment filed February 17, 2005).

By way of further example, according to the Molecular Biology of the Cell (Albert, 4th ed., Garland Science (2002)) (Exhibit C) (see Amendment filed February 17, 2005), one of the leading textbooks in the field of molecular biology:

> Extensive collaborative efforts are underway to generate comprehensive libraries of mutation in several model organisms including . . . the mouse. The ultimate goal in each case is to produce a collection of mutant strains in which every gene in the organism has either been systematically deleted, or altered such that it can be conditionally disrupted. Collections of this type will ***provide an invaluable tool for investigating gene function*** on a genomic scale.

(p. 543)(emphasis added).

According to Genes VII (Lewin, Oxford University Press (2000)) (Exhibit D) (see Amendment filed February 17, 2005), another well respected textbook in the field of genetics:

> The converse of the introduction of new genes is the ability to disrupt specific endogenous genes. Additional DNA can be introduced within a gene to prevent its expression and to generate a null allele. Breeding from an animal with a null allele can generate a homozygous "knockout", which has no active copy of the gene. This is ***a powerful method to investigate directly the importance and function of the gene.***

(p. 508)(emphasis added).

According to Joyner (Gene Targeting: *A Practical Approach*, preface, Oxford University Press 2000) (Exhibit E) (see Amendment filed February 17, 2005) "[g]ene targeting in ES cells offers a powerful approach to study gene function in a mammalian organism."

According to Matise et al. (*Production of Targeted Embryonic Stem Cell Clones* in Joyner, Gene Targeting: *A Practical Approach*, Oxford University Press 2000)(Exhibit F) (Amendment filed February 17, 2005):

> The discovery that cloned DNA introduced into tissue culture cells can undergo homologous recombination at specific chromosomal loci has revolutionized our ability to study gene function in cell culture and *in vivo*. . . . Thus, applying gene targeting technology to ES cells in culture affords researchers the opportunity to modify endogenous genes and ***study their function*** *in vivo*.

(p. 101)(emphasis added).

According to Crawley (What's Wrong With My Mouse *Behavioral Phenotyping of Transgenic and Knockout Mice*, Wiley-Liss 2000) (Exhibit G) (see Amendment filed February 17, 2005):

> Targeted gene mutation in mice represents a new technology that is revolutionizing biomedical research.

> Transgenic and knockout mutations provide an important means for understanding gene function, as well as for developing therapies for genetic diseases.

(p. 1, rear cover).

According to Doetschman (*Laboratory Animal Science* 49:137-143, 137 (1999)(Exhibit H)(emphasis added)(see Amendment filed February 17, 2005):

> [t]he conclusions will be that the ***knockout phenotypes do, in fact, provide accurate information concerning gene function,*** that we should let the unexpected phenotypes lead us to the specific cell, tissue, organ culture, and whole animal experiments that are relevant to the function of the genes in question, and that the absence of phenotype indicates that we have not discovered where or how to look for a phenotype.

Statements in the prior art, such as those provided above, may be used to establish that a person of ordinary skill in the art would immediately appreciate the asserted utility. *See In re Brana*, 51 F.3d at 1567 ("The prior art further supports the conclusion that one skilled in the art would be convinced of the asserted utility.") Accordingly, Appellant submits that in light of arguments of record a person of ordinary skill in the art would immediately appreciate one reason why the invention is useful: ***for studying the function of the PTP36 gene.***

In addition, Appellant notes that the National Institutes of Health (N.I.H.) has recently announced a three-year contract with the assignee of the instant application, Deltagen, Inc., to procure knockout mice. The Director of the N.I.H., Elias Zerhouni, M.D. has stated that the "decision to procure these knockout mouse lines and data and make them available to the research community will yield tremendous benefits, both in the short and long terms." *See* "Researchers to Gain Wider Access to Knockout Mice," available at <<http://www.genome.gov/17015131>> (Exhibit I). Thus, the NIH clearly believes that knockout mice

are useful, specifically including knockout mice generated by the assignee of the instant application.

(b) The Asserted Utility Of Using The PTP36 Knockout Mouse To Study The Function Of The PTP36 Gene Is Credible.

A "credible" assertion of utility is one that is

> [B]elievable to a person of ordinary skill in the art based on the totality of evidence and reasoning provided. An assertion is credible unless (A) the logic underlying the assertion is seriously flawed, or (B) the facts upon which the assertion is based are inconsistent with the logic underlying the assertion.

(MPEP 2107.02, III(B)(emphasis added).

In light of the arguments made above which show that the knockout mouse is the premier tool known in the art for determining gene function, the asserted utility of using the claimed mouse to study the function of the PTP36 gene would clearly be believable to a person of ordinary skill in the art. Therefore, the asserted use is credible.

(c) The Asserted Utility Of Using The PTP36 Knockout Mouse To Study The Function Of The PTP36 Gene Is A Specific Utility

Use of the PTP36 knockout mouse to study the function of the PTP36 gene is specific to this mouse. ***Only*** the PTP36 gene has been disrupted in the PTP36 knockout mouse. The ability to target specific genes using mouse ES cells is described below:

> One remarkable property of many mammalian cell lines, including ES cells, is the ability of transfected DNA to locate and recombine with its homologous chromosomal counterpart in the genome. This process, commonly referred to as gene targeting, is being used in many laboratories to generate ES cells and ultimately mouse lines in which

> specific loci have been modified. A common rationale to disrupt gene function by gene targeting in ES cells is to construct a vector which is designed to undergo a homologous genetic exchange with its chromosomal counterpart. These vectors are typically arranged so that they insert additional sequences, such as a positive selection marker, into the coding elements of the target, thereby ablating the function of the target gene.
>
> . . .
>
> Since gene targeting techniques are designed to generate specific recombinant alleles, features of the desired recombinant locus have been used to select in vitro for targeted clones and against random integration events, thereby enhancing the representation of targeted clones in selected populations.

(Bradley *et al.*, Biotechnology 10:534-538, 535-6 (1992) (cited by Examiner, paper no. 8)(emphasis added). Since each knockout mouse is specifically designed to disrupt a particular gene, the use of each mouse is specific for studying the function of ***that*** disrupted gene. Thus, only a PTP36 knockout mouse (as opposed to all other knockout mice) would be used to study the function of the PTP36 gene.

(d) Actual Use Of The PTP36 Knockout Mouse To Study Gene Function Constitutes A Real World Use

Determining the function of the PTP36 gene constitutes a real world use as demonstrated by Assignee Deltagen's phenotypic analyses of the claimed PTP36 knockout mouse and comparison with wild-type control mice. The claimed mice were ***actually*** generated. The claimed mice were ***successfully used in the asserted manner***—for determining the function of the PTP36 gene—by studying phenotypic changes in comparison with wild-type control mice. The claimed mice were subjected to an exhaustive battery of phenotypic screens including:

- physical examinations;
- necropsy, including body length, body weight, and organ weight measurements;

- histological examination of tissues and organs;
- bone marrow section evaluations;
- complete blood counts and differentials;
- clinical chemistry panels; and
- behavioral tests.

See Specification, page 48, lines 1-9.

Any statistically significant changes between the claimed mice and wild-type control mice are the result of the disruption of the PTP36 gene. *See* specification, page 18, line 21 through page 19, line 7. According to the specification (page 49, lines 14-23) (emphasis added):

> The homozygous female mice ***did not have demonstrable mammary gland tissue or had only a few mammary ducts on deeper levels.*** The homozygous female mice also contained the presence of keratin in the uterine horns; more specifically, keratinous debris was detected, but without concomitant squamous metaplasia, suggestive of passage of keratin from the vagina through a dilated cervix, indicating ***cervical relaxation.*** ***Increased anogenital distance*** in homozygous mutant females was also detected. The combination of lack of mammary growth and cervical relaxation are suggestive of a hormonal imbalance. Hormones having an effect on reproductive and/ or mammary tissues during development include estrogens, progesterones, growth hormone, thyroxine and insulin. ***The phenotypes observed are consistent with androgenization of the mutant female mice.***

Thus, by generating and studying the claimed mice, ***the Appellant has identified the function of the underlying protein encoding PTP36 gene,*** specifically the role of the PTP36 gene in the generation, *inter alia,* of mammary gland tissue. This stands in sharp contrast to the situation in *In re Fisher,* where the Federal Circuit's finding of lack of utility was based on the total failure of the appellant to show that the claimed ESTs had been successfully used in any of the asserted ways to determine the function of the underlying protein-encoding

gene. *See In re Fisher*, 421 F.3d at 1376 (holding that "precedent dictates that the [patent application] does not meet the utility requirement of § 101 because Fisher does not identify the function for the underlying protein-encoding genes.") In other words, the Appellant has studied the claimed invention "to the point of providing an immediate, well-defined, real-world benefit to the public meriting the grant of a patent." *Id.* Unlike the invention in *In re Fisher*, the claimed invention is not merely an "object upon which scientific research could be performed with no assurance that anything useful will be discovered in the end." *In re Fisher*, 421 F.3d at 1374. To the contrary: the claimed invention ***has*** provided the useful end result of defining the function of the underlying PTP36 gene. As one skilled in the art would immediately appreciate, ascribing functions to genes is the end goal of genomics research.

The Examiner has argued that using the claimed mouse to study the function of the PTP36 gene constitutes "further research," and that such "further research" is not a substantial utility. *See* Office Action mailed March 22, 2005, page 8. The Examiner characterizes the utility of using the claimed mouse to study the function of the PTP36 gene as a "scientific utility" rather than a "patentable utility." *See id.* In support of this proposition, the Examiner cites U.S. Pat. & Trademark Off., Manual of Patent Examining Procedure § 2107.01.I:

> Utilities that require or constitute carrying out further research to identify or reasonably confirm a "real world" context of use are not substantial utilities. ...[T]he following are examples of situations that require or constitute carrying out further research to identify or reasonably confirm a "real world" context of use and, therefore, do not define "substantial utilities":
> (A) Basic research such as studying the properties of the claimed product itself or the mechanisms in which the material is involved;

The language cited by the Examiner establishes that an invention has no substantial utility when "further research" is required to identify or confirm ***any*** utility. In other words, the cited language deals only with the situation where there is no known utility for a claimed invention, and where the only asserted use is studying the invention to determine a utility. This is clearly not the situation presented by the claimed PTP36 knockout mice. The Appellant has actually studied the claimed mouse and has discovered that it has observable phenotypes, such as a lack of mammary gland tissue. Thus, the Appellant has used the claimed invention in the manner asserted and has discovered that the PTP36 gene ***is*** involved in the generation of mammary gland tissue. Accordingly, the Appellant has established that the claimed mice ***can*** be used to study the role of the PTP36 gene in the development of mammary gland tissue.

To be sure, further research on the mouse will be required to ***refine*** the precise details of the molecular pathway through which PTP36 participates in the development of mammary gland tissue; however, the expectation of such further research is not inimical to a finding of utility. *See In re Brana*, 51 F.3d at 1568 ("Usefulness in patent law... necessarily includes the expectation of further research and development."). ***No*** further research is required to identify a utility for the PTP36 knockout mouse or to confirm that the PTP36 knockout mouse is useful for studying the role of PTP36 in the generation of mammary gland tissue. Put simply, there is a world of difference between using the mouse to determine ***any*** utility and using the mouse to determine the role of the PTP36 gene in, for example, the generation of mammary gland tissue.

The Examiner has argued that none of the described phenotypes "correlate with any known disease (either mouse or human)." Office Action, mailed March 22, 2005, page 8. As a result, the Examiner contends, "it is not

apparent to one of ordinary skilled [sic] in the art to immediately appreciate how to use the claimed mouse base [sic] on such characteristics." *Id.* As described in Section VII.E below, the phenotypes described are characteristic of hyperandrogenism (an excess of androgens) in females. Accordingly, the described phenotypes ***do*** correspond to a known disease in humans. In any event, it is entirely unnecessary for the described phenotypes to correlate to a known disease in order for the claimed knockout mouse to be useful in determining the function of the PTP36 gene. Useful information on the function of a gene can be obtained solely by studying phenotype(s), irrespective of the correlation of those phenotypes to a known disease. While the correlation of phenotypes to a disease provides ***additional*** useful information regarding the function of the gene, it is not a prerequisite—neither in the law nor from a practical, scientific viewpoint—for obtaining any useful functional information.

The Examiner has also argued that, in general, "knockout mice may not be capable of elucidating the function of the protein and may provide a clue to [the] pathway the protein being knocked out is involved in." Office Action mailed March 22, 2005, page 10 (emphasis omitted). In support of this proposition, the Examiner cites R.W. OLSEN & G.E. HOMANICS, *Function of GABA-A Receptors: Insights From Mutant and Knockout Mice, in* GABA IN THE NERVOUS SYSTEM: THE VIEW AT 50 YEARS 81-86 (D.L. Martin, and R.W. Olsen, eds., 200) as stating:

> Although gene targeting is often useful in delineating the contribution of a given gene product to phenotypic characteristics observed, ***some gene knockouts lead to embryonic or perinatal lethality, and others lead to no apparent phenotype.*** This can arise from the lack of any role for the gene in question in regard to the trait studies or from compensation by other gene products. Analysis of the compensation can yield valuable clues to the genetic pathway.

Id. at 82 (emphasis added).

The quote from Olsen & Homanics concerns only cases where there is early lethality or where there is no visible phenotype, *i.e.*, where the knockout mice appear identical to wild-type mice. This is clearly not the case for the claimed PTP36 knockout mice. Unlike the examples discussed by Olsen & Homanics, the claimed PTP36 knockout mice ***have*** phenotypes and do not exhibit perinatal or embryonic lethality. Moreover, Olsen & Homanics state that "gene targeting is often useful in delineating the contribution of a given gene product to phenotypic characteristics observed." *Id.* Accordingly, Olsen & Homanics actually confirm that the claimed PTP36 knockout mice have a substantial utility.

(e) Actual Use Of The Gene Function Data Derived From The Use Of The PTP36 Knockout Mouse Constitutes A Further Real World Use.

The claimed mouse has been extensively analyzed using the tests set forth in the Examples. Such tests reveal the function of the PTP36 gene in, *inter alia*, the generation of mammary gland tissue. This data has been incorporated into Appellant Deltagen's commercial database product, DeltaBase. Accordingly, the application provides:

> In one embodiment, the phenotype (or phenotypic change) associated with a disruption in the PTP36 gene is placed into or stored in a database. Preferably, the database includes: (i) genotypic data (*e.g.*, identification of the disrupted gene) and (ii) phenotypic data (*e.g.*, phenotype(s) resulting from the gene disruption) associated with the genotypic data. The database is preferably electronic. In addition, the database is preferably combined with a search tool so that the database is searchable.

Specification, page 14, lines 11-16. DeltaBase is the world's most comprehensive commercial library of in-depth functional information regarding *in vivo* mammalian gene function and their relevancy to small molecule drug

discovery. DeltaBase contains phenotypic information on approximately seven hundred and fifty target genes, including PTP36. DeltaBase offers the ability to not only further study the function of the PTP36 gene, but also provides a consistent analysis platform by allowing comparisons between data sets for all the gene knockouts contained in the database. *See* "Deltagen and Merck Enter Into DeltaBase License Agreement," available at <<http://www.sec.gov/Archives/edgar/data/1034072/000089843002000648/dex992.htm>>**(Exhibit J)**

The commercial success of the database is demonstrated by subscriptions to DeltaBase by at least three of the world's largest pharmaceutical companies, Merck, Pfizer and GSK. *See* Amendment filed June 22, 2005, page 7; ***see also*** **Declaration of Robert Driscoll, filed June 22, 2005 (Exhibit K).** To date, the sum of the subscription and related fees from just these three subscribers alone amounts to more than sixty million dollars ($60,000,000.00). This evidence of commercial success more than satisfies the practical utility requirement of 35 U.S.C. § 101 because "[p]roof of . . . utility is further supported when, as here, the inventions set forth in [the] claims . . . have on their merits been met with commercial success." *Raytheon Co. at* 959.

By creating a knockout mouse with a disruption in the PTP36 gene, and incorporating the data into a commercial product, and selling that data to pharmaceutical companies, Appellant has offered an "an immediate, well-defined, real-world benefit to the public." *In re Fisher*, 421 F.3d 1376. Unlike the Appellant in *In re Fisher*, Appellant has ***actually*** used the invention in the real world.

(f) Commercial Sale Of The PTP36 Knockout Mouse Constitutes A Further Real World Use.

The Examiner has argued that "[c]ommercial success is only considered as secondary evidence for overcoming a [Section] 103(a) rejection." Advisory Action mailed July 14, 2005. This statement of the law is incorrect. The Federal Circuit has clearly indicated that "commercial success may support the utility of an invention." *In re Fisher*, 421 F.3d at 1377-78 (citing *Raytheon Co.* 724 F.2d at 959 ("Proof of such utility is further supported when, as here, the inventions ... have on their merits been met with commercial success")).

At least one pharmaceutical company has licensed the presently claimed mouse. *See* **Amendment dated June 22, 2005, pages 7-8;** *see also* **Declaration of Robert Driscoll filed June 22, 2005 (Exhibit K).** The terms of the license agreement prevent the Appellant from revealing the name of the customer; however, the customer ranks among the top 10 pharmaceutical companies worldwide (based on sales). *See* Declaration of Robert Driscoll. The knockout mice provided by Deltagen are being used to study the function of the targeted gene and for human therapeutic drug development. *See* Driscoll Declaration, page 2, para. no. 4. According to the Federal Circuit in *In re Fisher*:

> [The Appellant] did not present any evidence showing that agricultural companies have purchased or even expressed any interest in the claimed ESTs. And, it is entirely unclear from the record whether such business entities ever will. Accordingly, while commercial success may support the utility of an invention, it does not do so in this case. See *Raytheon Co. v. Roper*, 724 F.2d 951, 959 (Fed. Cir. 1983)(stating that proof of a utility may be supported when a claimed invention meets with commercial success).

In re Fisher, 421 F.3d at 1377-78. In this case, Appellant ***has*** submitted evidence that the claimed invention has been purchased and delivered to at least

one large pharmaceutical company. Unlike the Appellant in *In re Fisher*, Appellant has presented evidence that the PTP36 knockout mouse ***has*** actually been used in the real world. Just as "[p]eople rarely, if ever, appropriate useless inventions," *Raytheon Co.* 724 F. 2d at 959, it is abundantly clear that large pharmaceutical companies rarely, if ever, ***purchase*** useless inventions. Thus, the delivery of the claimed PTP36 knockout mouse more than satisfies the practical utility requirement of Section 101 as it cannot be reasonably argued that a claimed invention which has been actually purchased and delivered to those skilled in the art has no "real world" use.

The Examiner argues commercial success is only established by "worldwide use" and "high revenue" from sales. *See* Advisory Action mailed July 14, 2005. The Examiner further argues that the sale of the claimed PTP36 knockout mouse to one pharmaceutical company does not amount to commercial success. *See id.* Appellant submits that commercial success must be determined on a market-by-market basis, rather than using an arbitrary yardstick such as "worldwide use." The market for PTP36 knockout mice—which, as described above, are used for validating drug targets—is comprised of only those few large pharmaceutical companies who have the resources to develop new pharmaceutical products. As is well known in the art, the cost of developing a new pharmaceutical is simply staggering: one recent economic study found that the average cost of research and development for a new drug is $802 million. *See* DiMiasi et al., Journal of Health Economics 22 (2003) 151–185 (Exhibit L). Therefore, the sale of the claimed PTP36 knockout mouse to at least one of the ten largest pharmaceutical companies in the world constitutes "commercial success."

2. The PTP36 Knockout Mice Have Further Well-Established Patentable Utility As They Are Useful For Validating The PTP36 Gene As A "Druggable" Target

PTP36 is a member of the protein tyrosine phosphatase family. Phosphatases are currently under intense scrutiny as targets for the development of drugs. *See* Huijsduijnen et al., Expert Opin. Ther. Targets, 2002, 6:624 (see Exhibit M). In one recent report, it was estimated that 4% of the ~3,000 genes in the "druggable genome" are protein phosphatases (the "druggable genome" is a subset of the human genome, and is comprised of those genes whose products are likely able to bind to drug-like molecules and are, thus, pharmacologically accessible targets). In other words, there are about 100 protein phosphatases that are capable of serving as drug targets. *See* Hopkins & Groom, Nature Reviews Drug Discovery 1:727 (2002) (Exhibit N).

The process of determining whether a given gene in the druggable genome is a worthwhile target for the development of a drug is termed "target validation." At the time of filing, and to this day, the "gold standard" for target validation is the knockout mouse. A knockout mouse mimicks the effect of treating a wild-type mouse with a drug that specifically inhibits the function of the target gene. By determining the phenotypic consequences of the disruption of the target gene, one can see the likely effect—and side effects—of such a drug ***without first incurring the massive expense of developing and testing a candidate drug.*** A recent article illustrates that the use of mouse knockouts in drug target validation was a well-established—and highly successful—utility at the time of filing:

> After a decade of using mouse knockouts, the data on their predictive power in drug discovery is ***irrefutable***. The top 100 selling drugs in 2001 are directed only to 29 drug targets, many with multiple agents addressing the same target. Of these 29 targets, 23 have been knocked

> out and in every case the knockout mouse was ***highly predictive as to the on-target effects and side effects of the associated drugs.***

See Arthur T. Sands, Industrializing Breakthrough Discovery, Current Drug Discovery, Aug. 2002, at 21.) (emphasis added) (Exhibit O).

Another phosphatase, PTP1B, was validated as a target using this approach. Specifically, since mice with disruptions in PTP1B are hyper-sensitive to insulin and resistant to diet-induced obesity, it was predicted that specific PTP1B inhibitors could be used to treat type 2 diabetes and obesity. *See* Elchebly et al., Science 283:1544 (1999) (Exhibit P). PTP1B inhibitors are currently the focus of considerable drug discovery efforts by major pharmaceutical companies.

The claimed mice mimic the effects of treating a wild-type mouse with a drug that specifically inhibits the activity of the PTP36 phosphatase. Thus, the claimed mice have "a significant and presently available benefit to the public," *In re Fisher*, 421 F.3d 1371, because they actually have been made and studied and ***actually have provided the public with the useful knowledge*** that the effects of a PTP36 inhibitor on a female would likely be a hormonal imbalance leading to androgenization. This information is critically useful for pharmaceutical companies in determining whether they should direct their research resources towards developing inhibitors of PTP36. In addition, this information is useful to pharmaceutical companies because it reveals the side effects of phosphatase inhibitors that are not totally specific for a targeted phosphatase. For example, if a phosphatase inhibitor drug targets the PTP1B phosphatase (a major target in diabetes and obesity), but also has an inhibitory effect on PTP36 *in vitro*, then the side effects of treatment with this PTP1B inhibitor can be predicted using information provided by the claimed mouse.

A recent review article has independently confirmed the utility of the claimed mice in validating phosphatases, specifically including PTP36, as druggable targets. Specifically, in **Huijsduijnen et al., Expert Opin. Ther. Targets, 2002, 6:624-643 (emphasis added) (Exhibit M)**, the authors state:

> Validating a PTP drug target is not a trivial task. Much more so than for other signalling components, predicting the effects of PTP inhibition from overexpression or substrate specificity studies alone has not been very successful. The history of PTP1B's discovery as a target clearly illustrates that purely deterministic approaches alone are not appropriate for PTPs.... An effective, systematic PTP target approach would instead focus on ***knockout animals*** and antisense or other inhibitors as tested in disease models.

In their discussion of the utility of knockout mice in validating PTPs as targets, Huijsduijnen et al. specifically refer to PCT Publication No. WO 02/45500—the PCT counterpart of the instant U.S. application—as teaching that PTP36 knockout female mice are androgenized with no mammary gland tissue. Thus, Huijsduijnen et al. independently confirm that the utility of using PTP36 knockout mice to validate PTP36 as a druggable target is well-established, substantial, and specific.

3. The Subject Matter Of Claims 28-32, 37, 47, 53-57 Has Further Patentable Utility As The PTP36 Knockout Mouse Is Useful As A Model For Studying Androgenization And For Identifying Therapeutic Agents For The Treatment Of Androgenization.

As described in Section VII.A.4.(a) *supra*, the claimed mice manifest androgenization, also known in the art as "hyperandrogenism." Hyperandrogenism is usually driven by excessive androgen (*i.e.*, male hormone) production by the ovaries, the adrenal glands, or both. It can result in hirsutism, acne, androgenic alopecia, oligo-ovulation, irregular menstrual bleeding, and

breast atrophy (reduction in mammary gland tissue). Hyperandrogenism may ultimately lead to life-threatening cardiovascular problems and metabolic disorders, such as type 2 diabetes mellitus. *See* American Association of Clinical Endocrinologists Medical Guidelines for Clinical Practice for the Diagnosis and Treatment of Hyperandrogenic Disorders, Endocrine Practice 7:121 (2001) (Exhibit Q). Approximately 7% of reproductive-aged women experience hyperandrogenism. *See* Azziz et al., J. Clin. Endocrinol. Metab. 2004, 89(2):453-62 (Exhibit R).

The claimed mouse therefore has a credible, specific, and substantial utility as a model of human disease, contrary to the Examiner's assertion. *See* Office Action mailed March 22, 2005 at page 5. Specifically, the claimed PTP36 knockout mouse has utility as a model of hyperandrogenism. As such, the claimed mice are useful for studying the pathogenesis of hyperandrogenism. The claimed mice—and tissues and cells isolated from the claimed mice—also have a credible, specific, and substantial utility as they can be used in methods for identifying drugs that treat hyperandrogenism. Again, this stands contrary to the Examiner's assertion that the claimed mice have no utility in the screening for therapeutic agents. *See* Office Action mailed March 22, 2005 at page 4. Methods for identifying such drugs are described at page 17, line 1 to page 18, line 15 of the specification.

4. The Subject Matter Of Claims 28-32, 37, 47, 53-57 Has Further Patentable Utility As The PTP36 Knockout Mouse Is Useful As A Model For Studying The Role Of PTP36 In Cancer Cell Metastasis

At the time of filing, it was well known in the art that PTP36 plays a role in regulating cell adhesion (see the Background of the Invention; see also Section

VII.C.4 *supra*). At the time of filing, it was also well known that cancer cells must disrupt the adhesive contacts they make with other cells in order to migrate from a primary tumor mass into the surrounding tissue *i.e.*, to metastasize. For example, see Cooper, The Cell: A Molecular Approach, Chapter 15 (2000) (Exhibit S) (emphases added):

> Cancer cells are also less stringently regulated than normal cells by cell-cell and cell-matrix interactions. ***Most cancer cells are less adhesive than normal cells,*** often as a result of reduced expression of cell surface adhesion molecules. For example, loss of E-cadherin, the principal adhesion molecule of epithelial cells, is important in the development of carcinomas (epithelial cancers). As a result of reduced expression of cell adhesion molecules, ***cancer cells are comparatively unrestrained by interactions with other cells and tissue components, contributing to the ability of malignant cells to invade and metastasize.*** The reduced adhesiveness of cancer cells also results in morphological and cytoskeletal alterations: Many tumor cells are rounder than normal, in part because they are less firmly attached to either the extracellular matrix or neighboring cells.

Because PTP36 was known to be involved in modulating cell adhesion, and because the modulation of cell adhesion was known to be involved in metastasis, one skilled in the art at the time of filing would immediately understand that the claimed PTP36 mice can be used to identify the role of the PTP36 gene in the growth and progression of tumors, specifically in the metastasis of cancer cells. Hence, the claimed mouse has a well-established, substantial, and specific utility in studying the role of PTP36 in cancer cell metastasis.

After the filing date of the instant application, the role of PTP36 mutations in the development of cancer was confirmed. *See* Wang et al, 2004, Science 304:1164 (Exhibit T) ; note that Wang et al refer to PTP36 using the synonymous name "PTPN14"). Specifically, it has been observed that inactivating mutations

in the PTP36 gene are observed in colorectal cancers. In addition, Wadham et al., Molecular Biology of the Cell 14:2520 (2003) (Exhibit U) have shown that inactivating mutations in PTP36 cause a decrease in cell-cell adhesion, with a concomitant increase in cell migration. These post-filing date references independently confirm the well-established, substantial, and specific utility of the claimed mice in the study of PTP36 in cancer cell metastasis.

5. The Subject Matter Of Claim 57 Has Further Patentable Utility As The PTP36 Knockout Mouse Is Useful For Studying Expression Of The PTP36 Gene.

The mouse of claim 57 further comprise a lacZ reporter gene. As explained in Section VII.A *supra,* the lacZ gene was inserted into the loci of the PTP36 gene and is under the control of the PTP36 endogenous promoter. Thus, lacZ is expressed where PTP36 is normally expressed. Expression may be studied by staining tissues and cells isolated from heterozygous knockout mice with X-gal and then observing the blue staining which is indicative of β-galactosidase expression. Expression analysis is therefore useful for analyzing ***where*** PTP36 is expressed.

According to the specification:

> *LacZ Reporter Gene Expression.* In general, tissues from 7-12 week old heterozygous mutant mice were analyzed for lacZ expression. Organs from heterozygous mutant mice were frozen, sectioned (10 μm), stained and analyzed for lacZ expression using X-Gal as a substrate for beta-galactosidase, followed by a Nuclear Fast Red counterstaining.
> In addition, for brain, wholemount staining was performed. The dissected brain was cut longitudinally, fixed and stained using X-Gal as the substrate for beta-galactosidase. The reaction was stopped by washing the brain in PBS and then fixed in PBS-buffered formaldehyde. Wild-type control tissues were also stained for lacZ expression to reveal any background or signals due to endogenous beta-galactosidase

> activity. The following tissues can show staining in the wild-type control sections and are therefore not suitable for X-gal staining: small and large intestines, stomach, vas deferens and epididymis. It has been previously reported that these organs contain high levels of endogenous beta-galactosidase activity.
>
> Striking lacZ (beta-galactosidase) expression was observed in blood vessels and adipocytes. LacZ expression was detectable in brain, eye, sciatic nerve, Harderian glands, thymus, spleen, aorta, heart, lung, liver, gall bladder, pancreas, kidney, urinary bladder, trachea, larynx, esophagus, thyroid gland, adrenal glands, salivary glands, tongue, skeletal muscle, skin, male and female reproductive systems.

Specification, page 50, lines 10-26. Using the claimed mice to study gene expression is specific since the lacZ gene is specifically inserted into the PTP36 gene, thereby falling under the control of the PTP36 gene promoter and thus revealing where the PTP36 gene is expressed.

Use of the claimed mice to study gene expression constitutes a substantial use as it represents a real world use. The use of heterozygous mice containing a reporter gene in studying gene expression is clearly recognized by those skilled in the art as a real world use:

> Null-reporter alleles should be created
> The project should generate alleles that are as uniform as possible, to allow efficient production and comparison of mouse phenotypes. The alleles should achieve a balance of utility, flexibility, throughput and cost. A null allele is an indispensable starting point for studying the function of every gene. Inserting a reporter gene (e.g., ***β-galactosidase*** or green fluorescent protein) allows a rapid assessment of which cell types normally support the expression of that gene.

Austin et al., Nature Genetics (2004) 36(9):921-24, 922) discussing the NIH's Knockout Mouse Project (emphasis in original; emphasis added)(Exhibit B) (see Amendment filed February 17, 2005).

Unlike the ESTs in *In re Fisher*, Appellant has ***actually*** used the invention in the asserted manner to generate useful information regarding the expression of the PTP36 gene. The claimed mice have provided "an immediate, well-defined, real-world benefit to the public," *In re Fisher*, 421 F.3d at 1376, because they ***actually have*** revealed where the PTP36 gene is expressed.

The Examiner argues that "studying the expression of a gene in which the function is unknown is not a substantial utility." Office Action dated March 22, 2005 at page 9. As shown above, studying the expression of a gene is a real-world use that is immediately appreciated as such by those skilled in the art. Moreover, contrary to the Examiner's assertion, the function of the PTP36 gene ***is*** known. *See* Section VII.C.4 *supra*. Using the claimed mouse to study the expression of a gene that has a known function is, contrary to the Examiner's assertion, a well-established and substantial utility. It is also a specific utility because only the claimed mouse having a LacZ gene inserted in the PTP36 gene can be used to study the expression of the PTP36 gene; mice having a LacZ gene inserted into another gene could not be used in this regard.

D. The Patent Application Enables One of Ordinary Skill in the Art to Use the Subject Matter of Claims 28-32, 37, 47, and 53-57 Because the Subject Matter of Claims 28-32, 37, 47, and 53-57 Has a Specific and Substantial Utility

The rejection of claims 28-32, 37, 47, and 53-57 under 35 U.S.C. § 112, first paragraph for lack of enablement is entirely contingent upon the rejection of claims 28-32, 37, 47, and 53-57 under 35 U.S.C. § 101 for lack of utility. Accordingly, because the claimed invention has a patentable utility, as demonstrated *supra*, the 35 U.S.C. § 112, first paragraph rejection should be withdrawn.

The Applicant also notes that the Examiner originally agreed that the invention was enabled. In the Office Action mailed June 3, 2003, the Examiner stated:

> The specification and the working examples provide sufficient guidance to use the invention of a female homozygous knockout mouse comprising a disruption [of] the PTP36 gene which result in no expression of the protein, wherein said mouse exhibits phenotypic features including androgenization, uterine abnormality, abnormal body or organ weight and abnormal physical feature as compared to wild type mice.

Id. at page 5.

Subsequently, the Examiner argued the specification failed to teach how to use "a PTP36 transgenic knockout mouse, wherein the homozygous knockout mouse exhibits phenotype of an uterine abnormality, hormonal imbalance, androgenization, increased body weight, increased organ weight, reduced or absent mammary tissue or increased anogenital distance." Office Action mailed April 27, 2004, page 3.

Appellant respectfully submits that the Examiner was correct the first time: the specification clearly teaches one skilled in the art how to use the claimed invention. Since the "how to use" prong of Section 112, first paragraph is satisfied, the utility requirement of Section 101 is also satisfied.

E. The Rejection of Claims 28-32, 37, 47, 53-54, and 57 under 35 U.S.C. § 112, first paragraph (new matter), Is Improper Because The Specification Reasonably Conveys That The Applicant Was In Possession Of The Claimed Subject Matter At the Time The Application Was Filed.

The Examiner rejected claims 28-32, 37, 47, 53-54, and 57 in the Office Action mailed March 22, 2005 because the specification allegedly does not disclose a knockout mouse comprising a null allele that comprises exogenous

DNA or a visible marker, as was recited in the claims pending on March 22, 2005. Without acquiescing in this rejection, the Appellant amended the claims pending in the Office Action mailed March 22, 2005 to remove the term "exogenous DNA" and to replace the term "visible marker" with "LacZ." *See* Supplemental Amendment mailed September 8, 2005. The Supplemental Amendment has been entered. Accordingly, it is believed that the rejection is overcome.

F. The Rejection of Claim 32 under 35 U.S.C. § 112, second paragraph, is Improper

In the Office Action of March 22, 2005, the Examiner advised the Appellant to add commas to the recitation "increased thymus weight increased liver weight relative to body weight and increased spleen weight relative to body weight" in claim 32. In the Supplemental Amendment mailed September 8, 2005, the Appellant made the requested changes. Accordingly, it is believed that the rejection is overcome.

VIII. CLAIMS APPENDIX.

The following claims are appealed and are presented as a clean copy. Claims canceled or withdrawn during prosecution are not listed.

28. A transgenic mouse whose genome comprises a null allele in the endogenous PTP36 gene.

29. The transgenic mouse of claim 54, wherein the female mouse exhibits, relative to a wild-type control mouse, a uterine abnormality comprising uterine dilation.

30. The transgenic mouse of claim 54, wherein the female mouse exhibits, relative to a wild-type control mouse, a uterine abnormality comprising keratin in the uterine horns.

31. The transgenic mouse of claim 54, wherein the female mouse exhibits, relative to a wild-type control mouse, a uterine abnormality comprising keratin in the uterine lumen.

32. The transgenic mouse of claim 54, wherein said mouse exhibits, relative to a wild-type control mouse, increased organ weight comprising at least one of the following: increased liver weight, increased spleen weight, increased thymus weight, increased liver weight relative to body weight, and increased spleen weight relative to body weight.

37. A cell or tissue isolated from the transgenic mouse of claim 28.

47. A method of producing the transgenic mouse of claim 28, the method comprising:

a. introducing a targeting construct capable of disrupting an endogenous PTP36 allele into a mouse embryonic stem cell;

b. selecting for the mouse embryonic stem cell that has undergone homologous recombination;

c. introducing the mouse embryonic stem cell selected for in step (b) into a blastocyst;

d. implanting the resulting blastocyst into a pseudopregnant mouse, wherein the resultant mouse gives birth to a chimeric mouse; and

e. breeding the chimeric mouse to produce the transgenic mouse.

53. The transgenic mouse of claim 28 wherein the mouse is heterozygous for said null allele.

54. The transgenic mouse of claim 28 wherein the mouse is homozygous for said null allele.

55. The transgenic mouse of claim 28 wherein said null allele comprises a gene encoding a selection marker.

56. The transgenic mouse of claim 55 wherein said gene is a neomycin resistance gene.

57. The transgenic mouse of claim 56 wherein said null allele further comprises a lacZ gene.

IX. EVIDENCE APPENDIX

Enclosed, please find copies of the following references:

Exhibit A

Background on Mouse as a Model Organism
<<http://www.genome.gov/pfv.cfm?pageid=10005834>>

Exhibit B

Austin et al., Nature Genetics (2004) 36(9):921-24, 921

Exhibit C

Alberts, Molecular Biology of the Cell, p.543 (4th Ed. Garland Science (2002))

Exhibit D

Genes VII, p.508 (Lewin, Oxford University Press (2000))

Exhibit E

Joyner, Gene Targeting: A Practical Approach, preface, Oxford University Press (2000)

Exhibit F

Matise et al., Production of Targeted Embryonic Stem Cell Clones, p.101 in Joyner, Gene Targeting: A Practical Approach, Oxford University Press 2000)

Exhibit G

Crawley, What's Wrong With My Mouse Behavioral Phenotyping of Transgenic and Knockout Mice, p.1, rear cover, Wiley-Liss 2000)

Exhibit H

Doetschman, Laboratory Animal Science 49:137-143, 137 (1999)

Exhibit I

Researchers to Gain Wider Access to Knockout Mice

<<http://www.genome.gov/17015131>>

Exhibit J

Bradley et al., Biotechnology 10:534-538, 535-6 (1992)

Exhibit K

Deltagen and Merck Enter Into DeltaBase License Agreement
<<http://www.sec.gov/Archives/edgar/data/1034072/000089843002000648/dex992.htm>>

Exhibit L

Declaration of Robert Driscoll, filed June 22, 2005

Exhibit M

DiMiasi et al., Journal of Health Economics 22 (2003) 151–185

Exhibit N

Huijsduijnen et al., Expert Opin. Ther. Targets, 2002, 6:624

Exhibit O

Hopkins & Groom, Nature Reviews Drug Discovery 1:727 (2002)

Exhibit P

Arthur T. Sands, Current Drug Discovery, Aug. 2002, at 21.

Exhibit Q

Elchebly et al., Science 283:1544 (1999)

Exhibit R

American Association of Clinical Endocrinologists Medical Guidelines for Clinical Practice for the Diagnosis and Treatment of Hyperandrogenic Disorders, Endocrine Practice 7:121 (2001)

Exhibit S

Azziz et al., J. Clin. Endocrinol. Metab. 2004, 89(2):453-62

Exhibit T

Cooper, The Cell: A Molecular Approach, Chapter 15 (2000)

Exhibit U

Wang et al, 2004, Science 304:1164

Exhibit V

Wadham et al., Molecular Biology of the Cell 14:2520 (2003)

X. CLOSING REMARKS

For the foregoing reasons, Appellant submits that the rejection of claims 28-32, 37, 47 and 53-57 under 35 U.S.C. § 101 for lack of utility is improper; that the rejection of claims 28-32, 37, 47 and 53-57 under 35 U.S.C. § 112, first paragraph, for lack of enablement is improper; that the rejection of claims 28-32, 37, 47, 53-54, and 57 under 35 U.S.C. § 112, first paragraph, for lack of written description is improper; and that the rejection of claim 32 for indefiniteness under 35 U.S.C. § 112, second paragraph, is improper.

Please charge the Appeal Brief fee of $250 for a small entity to Deposit Account No. 502775. It is believed that no other fees are due with this Appeal Brief. If this is in error, please charge any additional fees to Deposit Account No. 502775.

This constitutes a request for any needed extension of time under 37 C.F.R. § 1.136(a) and an authorization to charge all fees therefore to deposit account No. 502775 if not otherwise specifically requested.

Respectfully submitted,

11/16-05
Date

John E. Burke, Reg. No. 35,836
Greenberg Traurig LLP
1200 17th Street, Suite 2400
Denver, CO 80202
(303) 685-7411
(720) 904-6111 (fax)

EXHIBIT A

Background on Mouse as a Model Organism
<<http://www.genome.gov/pfv.cfm?pageid=10005834>>

URL of this page: http://www.genome.gov/10005834

Background on Mouse as a Model Organism

Over the past century, the mouse has developed into the premier mammalian model system for genetic research. Scientists from a wide range of biomedical fields have gravitated to the mouse because of its close genetic and physiological similarities to humans, as well as the ease with which its genome can be manipulated and analyzed.

Although yeasts, worms and flies are excellent models for studying the cell cycle and many developmental processes, mice are far better tools for probing the immune, endocrine, nervous, cardiovascular, skeletal and other complex physiological systems that mammals share. Like humans and many other mammals, mice naturally develop diseases that affect these systems, including cancer, atherosclerosis, hypertension, diabetes, osteoporosis and glaucoma. In addition, certain diseases that afflict humans but normally do not strike mice, such as cystic fibrosis and Alzheimer's, can be induced by manipulating the mouse genome and environment. Adding to the mouse's appeal as a model for biomedical research is the animal's relatively low cost of maintenance and its ability to quickly multiply, reproducing as often as every nine weeks.

Mouse models currently available for genetic research include thousands of unique inbred strains and genetically engineered mutants. There are mice prone to different cancers, diabetes, obesity, blindness, Lou Gehrig's disease, Huntington's disease, anxiety, aggressive behavior, alcoholism and even drug addiction. Immunodeficient mice can also be used as hosts to grow both normal and diseased human tissue, a boon for cancer and AIDS research.

In the early days of biomedical research, scientists developed mouse models by selecting and breeding mice to produce offspring with the desired traits. Researchers also learned to produce useful, new models of genetic disease quickly and in large numbers by exposing mice to DNA-damaging chemicals, a process known as chemical mutagenesis.

In recent decades, researchers have utilized an array of innovative genetic technologies to produce custom-made mouse models for a wide array of specific diseases, as well as to study the function of targeted genes. One of the most important advances has been the ability to create transgenic mice, in which a new gene is inserted into the animal's germline. Even more powerful approaches, dependent on homologous recombination, have permitted the development of tools to "knock out" genes, which involves replacing existing genes with altered versions; or to "knock in" genes, which involves altering a mouse gene in its natural location. To preserve these extremely valuable strains of mice and to assist in the propagation of strains with poor reproduction, researchers have taken advantage of state-of-the-art reproductive technologies, including cryopreservation of embryos, *in vitro* fertilization and ovary transplantation.

The Jackson Laboratory, a publicly supported national repository for mouse models in Bar Harbor, Maine, has played a crucial role in the development of the mouse into the leading model for biomedical research. Established in 1929, the non-profit center pioneered the use of inbred laboratory mice to uncover the genetic basis of human development and disease. In fact, the famous "Black 6" or C57BL/6J mouse strain whose genome is the focus of the landmark sequencing effort was developed in the early 1920s by The Jackson Laboratory

founder Clarence Cook Little.

Today, researchers at The Jackson Laboratory pursue projects in areas that include cancer, development and aging, immune system and blood disorders, neurological and sensory disorders, and metabolic diseases. Informatics researchers work with the public sequencing consortium to curate and integrate the sequenced mouse genome data with the wealth of biological knowledge collected in Jackson's Mouse Genome Informatics resource.

In addition, The Jackson Laboratory distributes 2,700 different strains and stocks as breeding mice, frozen embryos or DNA samples. In FY 2002 alone, the lab supplied approximately 2 million mice to the international scientific community.

Listed below is a sampling of mouse models developed and/or distributed by The Jackson Laboratory, along with brief descriptions of the human diseases they are helping scientists to understand:

- Down Syndrome - One of the most common genetic birth defects in humans, occurring once in every 800 to 1,000 live births, Down syndrome results from an extra copy of chromosome 21, an abnormality known as trisomy. The Ts65Dn mouse, developed at The Jackson Laboratory, mimics trisomy 21 and exhibits many of the behavioral, learning, and physiological defects associated with the syndrome in humans, including mental deficits, small size, obesity, hydrocephalus and thymic defects. This model represents the latest and best improvement of Down syndrome models to facilitate research into the human condition.

- Cystic Fibrosis (CF) - The *Cftr* knockout mouse has helped advance research into cystic fibrosis, the most common fatal genetic disease in the United States today, occurring in approximately one of every 3,300 live births. Scientists now know that CF is caused by a small defect in the gene that manufactures CFTR, a protein that regulates the passage of salts and water in and out of cells. Studies with the *Cftr* knockout have shown that the disease results from a failure to clear certain bacteria from the lung, which leads to mucus retention and subsequent lung disease. These mice have become models for developing new approaches to correct the CF defect and cure the disease.

- Cancer - The p53 knockout mouse has a disabled *Trp53* tumor suppressor gene that makes it highly susceptible to various cancers, including lymphomas and osteosarcomas. The mouse has emerged as an important model for human Li-Fraumeni syndrome, a form of familial breast cancer.

- Glaucoma - The DBA/2J mouse exhibits many of the symptoms that are often associated with human glaucoma, including elevated intraocular pressure. Glaucoma is a debilitating eye disease that is the second leading cause of blindness in the United States.

- Type 1 Diabetes - This autoimmune disease, also known as Juvenile Diabetes, or Insulin Dependent Diabetes Mellitus (IDDM), accounts for up to 10 percent of diabetes cases. Non-obese Diabetic (NOD) mice are enabling researchers to identify IDDM susceptibility genes and disease mechanisms.

- Type 2 Diabetes - A metabolic disorder also called Non-Insulin Dependent Diabetes Mellitus (NIDDM), this is the most common form of diabetes and occurs primarily after age 40. The leading mouse models for NIDDM and obesity research were all developed at The Jackson Laboratory: Cpe^{fat}, Lep^{ob}, $Lepr^{db}$ and *tub.*

- Epilepsy - The "slow-wave epilepsy," or *swe,* mouse is the only model to exhibit both of

the two major forms of epilepsy: petit mal (absence) and grand mal (convulsive). It shows particular promise for research into absence seizures, which occur most often in children.

- Heart Disease - Elevated blood cholesterol levels and plaque buildup in arteries within three months of birth (even on a low-fat diet) are characteristics of several experimental models for human atherosclerosis: the *Apoe* knockout mouse and C57BL/6J.

- Muscular Dystrophy - The *Dmd* ^{mdx} mouse is a model for Duchenne Muscular Dystrophy, a rare neuromuscular disorder in young males that is inherited as an X-linked recessive trait and results in progressive muscle degeneration.

- Ovarian Tumors - The SWR and SWXJ mouse models provide excellent research platforms for studying the genetic basis of ovarian granulosa cell tumors, a common and very serious form of malignant ovarian tumor in young girls and post-menopausal women.

Contact:
Geoff Spencer
NHGRI
Phone: (301) 402-0911

Top of page

Last Reviewed: September 2005

EXHIBIT B

Austin et al., Nature Genetics (2004) 36(9):921-24, 921

The Knockout Mouse Project

Mouse knockout technology provides a powerful means of elucidating gene function *in vivo*, and a publicly available genome-wide collection of mouse knockouts would be significantly enabling for biomedical discovery. To date, published knockouts exist for only about 10% of mouse genes. Furthermore, many of these are limited in utility because they have not been made or phenotyped in standardized ways, and many are not freely available to researchers. It is time to harness new technologies and efficiencies of production to mount a high-throughput international effort to produce and phenotype knockouts for all mouse genes, and place these resources into the public domain.

Now that the human and mouse genome sequences are known[1–3], attention has turned to elucidating gene function and identifying gene products that might have therapeutic value. The laboratory mouse (*Mus musculus*) has had a prominent role in the study of human disease mechanisms throughout the rich, 100-year history of classical mouse genetics, exemplified by the lessons learned from naturally occurring mutants such as agouti[4], reeler[5] and obese[6]. The large-scale production and analysis of induced genetic mutations in worms, flies, zebrafish and mice have greatly accelerated the understanding of gene function in these organisms. Among the model organisms, the mouse offers particular advantages for the study of human biology and disease: (i) the mouse is a mammal, and its development, body plan, physiology, behavior and diseases have much in common with those of humans; (ii) almost all (99%) mouse genes have homologs in humans; and (iii) the mouse genome supports targeted mutagenesis in specific genes by homologous recombination in embryonic stem (ES) cells, allowing genes to be altered efficiently and precisely.

The ability to disrupt, or knock out, a specific gene in ES cells and mice was developed in the late 1980s (ref. 7), and the use of knockout mice has led to many insights into human biology and disease[8–11]. Current technology also permits insertion of 'reporter' genes into the knocked-out gene, which can then be used to determine the temporal and spatial expression pattern of the knocked-out gene in mouse tissues. Such marking of cells by a reporter gene facilitates the identification of new cell types according to their gene expression patterns and allows further characterization of marked tissues and single cells.

Appreciation of the power of mouse genetics to inform the study of mammalian physiology and disease, coupled with the advent of the mouse genome sequence and the ease of producing mutated alleles, has catalyzed public and private sector initiatives to produce mouse mutants on a large scale, with the goal of eventually knocking out a substantial portion of the mouse genome[12,13]. Large-scale, publicly funded gene-trap programs have been initiated in several countries, with the International Gene Trap Consortium coordinating certain efforts and resources[14–17].

Despite these efforts, the total number of knockout mice described in the literature is relatively modest, corresponding to only ~10% of the ~25,000 mouse genes. The curated Mouse Knockout & Mutation Database lists 2,669 unique genes (C. Rathbone, personal communication), the curated Mouse Genome Database lists 2,847 unique genes, and an analysis at Lexicon Genetics identified 2,492 unique genes (B.Z., unpublished data). Most of these knockouts are not readily available to scientists who may want to use them in their research; for example, only 415 unique genes are represented as targeted mutations in the Jackson Laboratory's Induced Mutant Resource database (S. Rockwood, personal communication).

The converging interests of multiple members of the genomics community led to a meeting to discuss the advisability and feasibility of a dedicated project to produce knockout alleles for all mouse genes and place them into the public domain. The meeting took place from 30 September to 1 October 2003 at the Banbury Conference Center at Cold Spring Harbor Laboratory. The attendees of the meeting are the authors of this paper.

Is a systematic project warranted?

A coordinated project to systematically knock out all mouse genes is likely to be of enormous benefit to the research community, given the demonstrated power of knockout mice to elucidate gene function, the frequency of unpredicted phenotypes in knockout mice, the potential economies of scale in an organized and carefully planned project, and the high cost and lack of availability of knockout mice being made in current efforts. Moreover, implementing such a systematic and comprehensive plan will greatly accelerate the translation of genome sequences into biological insights. Knockout ES cells and mice currently available from the public and private sectors should be incorporated into the genome-wide initiative as much as possible, although some may be need to be produced again if they were made with suboptimal methods (*e.g.*, not including a marker) or if their use is restricted by intellectual property or other constraints. The advantages of such a systematic and coordinated effort include efficient production with reduced costs; uniform use of knockout methods, allowing for more comparability between knockout mice; and ready access to mice, their derivatives and data to all researchers without encumbrance. Solutions to the logistical, organizational and informatics issues associated with producing, characterizing and distributing such a large number of

*The Comprehensive Knockout Mouse Project Consortium**
**Authors and their affiliations are listed at the end of the paper.*

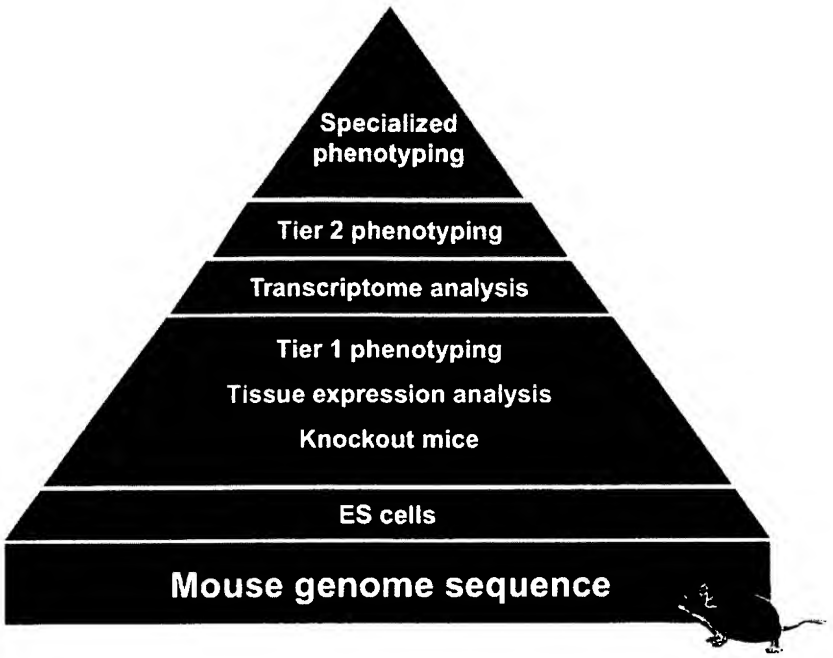


Figure 1 Structure of resource production in the proposed KOMP. Using the mouse genome sequence as a foundation, knockout alleles in ES cells will be produced for all genes. A subset of ES cell knockouts will be used each year to produce knockout mice, determine the expression pattern of the targeted gene in a variety of tissues and carry out screening-level (Tier 1) phenotyping. In a subset of mouse lines, transcriptome analysis and more detailed system-specific (Tier 2) phenotyping will be done. Finally, specialized phenotyping will be done on a smaller number of mouse lines with particularly interesting phenotypes. All stages will occur within the purview of the KOMP except for the specialized phenotyping, which will occur in individual laboratories with particular expertise.

mice will draw from the experience of related projects in the private sector and in academia, which have made or phenotyped hundreds of knockout mice using a variety of techniques. Lessons learned from these projects include the need for redundancy at each step to mitigate pipeline bottlenecks and the need for robust informatics systems to track the production, analysis, maintenance and distribution of thousands of targeting constructs, ES cells and mice.

Null-reporter alleles should be created

The project should generate alleles that are as uniform as possible, to allow efficient production and comparison of mouse phenotypes. The alleles should achieve a balance of utility, flexibility, throughput and cost. A null allele is an indispensable starting point for studying the function of every gene. Inserting a reporter gene (*e.g.*, β-galactosidase or green fluorescent protein) allows a rapid assessment of which cell types normally support the expression of that gene. Therefore, we propose to produce a null-reporter allele for each gene. Making each mutation conditional in nature by adding *cis*-elements (*e.g.*, *lox*P or FRT sites) would be desirable, but we do not advocate this as part of the mutagenesis strategy unless the technological limitations currently associated with generating conditional targeted mutations on a large scale and in a cost-effective manner can be overcome.

A combination of methods should be used

Various methods can be used to create mutated alleles, including gene targeting, gene trapping and RNA interference. Advantages of conventional gene targeting include flexibility in design of alleles, lack of limitation to integration hot spots, reliability for producing complete loss-of-function alleles, ability to produce reporter knock-ins and conditional alleles, and ability to target splice variants and alternative promoters. BAC-based targeting has the potential advantages of higher recombination efficiencies and flexibility for producing complex mutated alleles[18]. Gene trapping is rapid, is cost-effective and produces a large variety of insertional mutations throughout the genome but can be somewhat less flexible[17,19–21]. There is uncertainty regarding the percentage of gene traps that produce a true null allele and the fraction of the genome that can ultimately be covered by gene-trap mutations. Trapping is not entirely random but shows preference for larger transcription units and genes more highly expressed in ES cells. In recent studies, gene trapping was estimated to potentially produce null alleles for 50–60% of all genes, perhaps more if a variety of gene-trap vectors with different insertion characteristics is used[17,21]. RNA interference offers enormous promise for analysis of gene function in mice[22] but is not yet sufficiently developed for large-scale production of gene modifications capable of reliably producing true null alleles. Both gene-targeting and gene-trapping methods are suitable for producing large numbers of knockout alleles, and, given their complementary advantages, a combination of these methods should be used to produce the genome-wide collection of null-reporter alleles most efficiently.

What should the deliverables be?

A genome-wide knockout mouse project could deliver to the research community a trove of valuable reagents and data, including targeting and trapping constructs and vectors, mutant ES cell lines, live mice, frozen sperm, frozen embryos, phenotypic data at a variety of levels and detail, and a database with data visualization and mining tools. At a minimum, we believe that a comprehensive genome-wide resource of mutant ES cell lines from an inbred strain, each with a different gene knocked out, should be produced and made available to the community. Choosing an inbred line (129/SvEvTac or C57BL/6J), and evaluating the alternative of using F_1 ES cells and tetraploid aggregation to provide potential time savings, merits additional scientific review and discussion[23,24]. ES cells should be converted into mice at a rate consistent with project funding and the ability of the worldwide scientific community to analyze them. Although the value and cost-effectiveness of systematically characterizing the mice is a matter of debate, a limited set of broad and cost-effective screens, probably including assessment of developmental lethality, physical examination, basic blood tests, and histochemical analysis of reporter gene expression, would be useful. More detailed phenotyping, based on findings from the initial screen or existing knowledge of the gene's function, could be done at specialized centers. All ES cell clones and mice (as frozen embryos or sperm) should be available to any researcher at minimal cost, and all mouse phenotyping and reporter expression data should be deposited into a public database.

In determining how to implement the project, utility to the research community should be the standard for judging value. Each step after ES cell generation (*e.g.*, mouse creation, breeding, expression analysis, phenotyping) will make the resource useful to more researchers but will also increase costs and scientific complexity. We therefore advocate a 'pyramid' structure for the project (**Fig. 1**). At the base of the pyramid is the genome-wide collection of mutant ES cells for every mouse gene. Over time, a subset of these mutant ES cells should be made into mice and characterized with an initial phenotype screen (Tier 1; **Fig. 1**) and analysis of tissue reporter-gene expression. A subset of these lines should be profiled by microarray analysis, and a subset of these profiled by system-specific (Tier 2) phenotyping, based on the results of the Tier 1 phenotyping, array studies, existing knowledge of the gene's function and the gene's tissue expression pattern. With time, the upper tiers of the pyramid will be filled out, eventually transforming the pyramid into a cube, with information of all types available for all genes.

This project will require the resolution of numerous intellectual property claims involving the production and use of knockout mice. To deal with the existing patents that cover the technologies and processes involved in the production of mutant mice, we suggest that a 'patent pool', such as that used in the semiconductor industry[25], should be generated. Several individuals who represent entities that control patents on mouse knockout technologies are authors on this paper, and they agree with this approach. We also agree that any mutant ES cells or mice produced should be placed immediately in the public domain.

Mechanisms and costs

ES cell production. Automated knockout construct and ES cell production should be carried out in coordinated centers to ensure efficiency and uniformity. We estimate that most known mouse genes could be knocked out in ES cells within 5 years, using a combination of gene-trapping and gene-targeting techniques. Gene trapping can produce a large number of mutated alleles quickly, but its progress should be monitored closely to determine when its yield of new genes diminishes[17] and, therefore, when targeting should be increasingly relied on. As large-scale trapping projects have already defined gene classes that probably cannot be knocked out by trapping (*e.g.*, single-exon GPCRs, genes that are not expressed in ES cells), we propose that targeting begin on those classes immediately. All ES cells should be made available to the research community, because this collection itself would be a valuable resource. Efforts in the public and private sectors have already knocked out many genes in ES cells, and, to the degree that the alleles produced fit the prescribed characteristics (*i.e.*, null alleles with a reporter) and are available, every effort should be made to incorporate these into the planned public resource. Costs for generating this part of the resource were estimated at between $9–11 million/year for five years (these and all subsequent figures are direct costs).

Mouse production. The subset of ES cells made into mice each year should be chosen by a peer-review process. Central facilities for high-efficiency mouse production, genotyping, breeding, maintenance and archiving should be funded, to take advantage of efficiencies of scale in mouse creation and distribution. Researchers could apply to produce groups of mice outside the centers, as long as they meet the cost specifications of the program. All mice should be made available immediately to researchers as frozen embryos or sperm, for nominal distribution cost. An initial target of 500 new mouse lines per year would double the current rate at which new genes are knocked out in the public sector; we feel that this rate is within the capacity of the biomedical research community worldwide to absorb and analyze. We estimated the initial cost of this level of mouse production to be $12.5–15 million per year.

Reporter tissue expression analysis. Approximately 30 tissues from adult and developmental stages should be sampled to cover the main organ systems. Analysis methods should be customized to the organ system and marker, and a searchable database of the sites of gene expression, and the images showing them, should be produced. Centers to carry out these analyses and data curation should be selected by peer review. We estimated the cost of this component for 500 mouse lines to be $2.5–5 million per year, depending on how much tissue sectioning and cell-level analysis is done.

Phenotyping. Tier 1 phenotyping should be a low-cost screen for clear phenotypes and should be done on all mouse lines produced. Tier 1 should include home-cage observation, physical examination, blood hematological and chemistry profiles, and skeletal radiographs. The centers producing the mice should carry out the Tier 1 analyses, at an estimated cost of $2.5 million per year for 500 lines. Selected lines, chosen on the basis of findings from Tier 1 phenotyping, tissue expression patterns, microarray data and the scientific literature, should undergo more detailed and system-focused Tier 2 phenotyping. Tier 2 phenotyping should be done in specialized phenotyping centers, akin to those already in operation for phenotyping of mice produced by ENU mutagenesis. All Tier 1 and Tier 2 phenotyping should be done on a uniform genetic background by dedicated groups of individuals in single locations, to facilitate consistency and cross-comparison of results among different mouse lines. All Tier 1 and Tier 2 phenotyping results should be deposited into a central project database freely accessible to the research community. More detailed and specialized phenotyping could be done by individual researchers in their own laboratories; deposition of this more detailed phenotype data would be encouraged.

Transcriptome analysis. Transcriptome profiling of tissues from each knockout line, collected in a uniform way across all mice and tissues and placed into a searchable relational database, would add substantially to the scientific value of the project, though it would also add considerably to its cost. Transcriptome analysis should therefore be done on a subset of mice, chosen by peer review. We estimate that, with the best currently available array technology, an analysis of ten tissues would cost ~$18,000 per line.

Conclusions

This project, tentatively named the Knockout Mouse Project (KOMP), will be a crucial step in harnessing the power of the genome to drive biomedical discovery. By creating a publicly available resource of knockout mice and phenotypic data, KOMP will knock down barriers for biologists to use mouse genetics in their research. The scientific consensus that we achieved—that a dedicated project should be undertaken to produce mutant mice for all genes and place them into the public domain—is important but is only the beginning. Implementation of these recommendations will require additional input from the greater scientific community, including those responsible for programmatic direction and financial support of biomedical research in the public and private sectors. This ambitious and historic initiative must be carried out as a collaborative effort of the worldwide scientific community, so that all can contribute their skills, and all can benefit. International discussions among scientific and programmatic staffs since the Banbury meeting at Cold Spring Harbor, in both the public and private sectors, have shown that there is great enthusiasm and commitment to this vision. The next step for KOMP will be to move this visionary plan from conceptualization to implementation, with an urgency befitting the benefits it will bring to science and medicine.

URLs. The curated Mouse Knockout & Mutation Database is available at http://research.bmn.com/mkmd/. The curated Mouse Genome Database is available at http://www.informatics.jax.org/. *Patent pools: A solution to the problem of access in biotechnology patents?* is available at http://www.uspto.gov/web/offices/pac/dapp/opla/patentpool.pdf.

1. International Human Genome Sequencing Consortium. *Nature* **409**, 860–921 (2001).
2. Venter, J.C. *et al. Science* **291**, 1304–1351 (2001).
3. Mouse Genome Sequencing Consortium. *Nature* **420**, 520–562 (2002).
4. Bultman, S.J., Michaud, E.J. & Woychik, R.P. *Cell* **71**, 1195–1204 (1992).
5. D'Arcangelo, G. *et al. Nature* **374**, 719–723 (1995).
6. Zhang, Y. *et al. Nature* **372**, 425–432 (1994).
7. Goldstein, J.L. *Nat. Med.* **7**, 1079–1080 (2001).
8. D'Orleans-Juste, P., Honore, J.C., Carrier, E. & Labonte, J. *Curr. Opin. Pharmacol.* **3**, 181–185 (2003).
9. Horton, W.A. *Lancet* **362**, 560–569 (2003).
10. Wallace, D.C. *Am. J. Med. Genet.* **106**, 71–93 (2001).
11. Chen, R.Z., Akbarian, S., Tudor, M. & Jaenisch, R. *Nat. Genet.* **27**, 327–331 (2001).
12. Zambrowicz, B.P. *et al. Nature* **392**, 608–611 (1998).
13. Nadeau, J.H. *et al. Science* **291**, 1251–1255 (2001).
14. Wiles, M.V. *et al. Nat. Genet.* **24**, 13–14 (2000).
15. Stryke, D. *et al. Nucleic Acids Res.* **31**, 278–281 (2003).
16. Hansen, J. *et al. Proc. Natl. Acad. Sci. USA* **100**, 9918–9922 (2003).
17. Skarnes, W.C. *et al. Nat. Genet.* **36**, 543–544 (2004).
18. Valenzuela, D.M. *et al. Nat. Biotechnol.* **21**, 652–629 (2003).
19. Chen, W.V., Delrow, J., Corrin, P.D., Frazier, J.P. & Soriano, P. *Nat. Genet.* **36**, 304–312 (2004).
20. Stanford, W.L., Cohn, J.B. & Cordes, S.P. *Nat. Rev. Genet.* **2**, 756–768 (2001).
21. Zambrowicz, B.P. *et al. Proc. Natl. Acad. Sci. USA* **100**, 14109–14114 (2003).
22. Kunath, T. *et al. Nat. Biotechnol.* 21, 559–561 (2003).
23. Seong, E., Saunders, T.L., Stewart, C.L. & Burmeister, M. *Trends Genet.* **20**, 59–62 (2004).
24. Eggan, K. *et al. Nat. Biotechnol.* **20**, 455–459 (2002).
25. Clark, J., Piccolo, J., Stanton, B. & Tyson, K. Patent pools: A solution to the problem of access in biotechnology patents? (US Patent and Trademark Office, 2000).

Christopher P Austin[1], James F Battey[2], Allan Bradley[3], Maja Bućan[4], Mario Capecchi[5], Francis S Collins[6], William F Dove[7], Geoffrey Duyk[8], Susan Dymecki[9], Janan T Eppig[10], Franziska B Grieder[11], Nathaniel Heintz[12], Geoff Hicks[13], Thomas R Insel[14], Alexandra Joyner[15], Beverly H Koller[16], K C Kent Lloyd[17], Terry Magnuson[18], Mark W Moore[19], Andras Nagy[20], Jonathan D Pollock[21], Allen D Roses[22], Arthur T Sands[23], Brian Seed[24], William C Skarnes[25], Jay Snoddy[26], Philippe Soriano[27], David J Stewart[28], Francis Stewart[29], Bruce Stillman[28], Harold Varmus[30], Lyuba Varticovski[31], Inder M Verma[32], Thomas F Vogt[33], Harald von Melchner[34], Jan Witkowski[35], Richard P Woychik[36], Wolfgang Wurst[37], George D Yancopoulos[38], Stephen G Young[39] & Brian Zambrowicz[40]

[1]National Human Genome Research Institute, National Institutes of Health, Building 31, Room 4B09, 31 Center Drive, Bethesda, Maryland 20892, USA. [2]National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Building 31, Room 3C02, Bethesda, Maryland 20892, USA. [3]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [4]Department of Genetics, University of Pennsylvania, 111 Clinical Research Building, 415 Curie Boulevard, Philadelphia, Pennsylvania 19104-6145, USA. [5]University of Utah, Eccles Institute of Human Genetics, Suite 5400, Salt Lake City, Utah 85112, USA. [6]National Human Genome Research Institute, National Institutes of Health, Building 31, Room 4B09, 31 Center Drive, Bethesda, Maryland 20892, USA. [7]McArdle Laboratory for Cancer Research, University of Wisconsin - Madison, 1400 University Avenue, Madison, Wisconsin 53706-1599, USA. [8]TPG Ventures, 345 California Street, Suite 2600, San Francisco, California 94104, USA. [9]Harvard Medical School, Department of Genetics, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA. [10]The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609-1500, USA. [11]National Center for Research Resources, National Institutes of Health, 1 Democracy Plaza, 6701 Democracy Boulevard, Bethesda, Maryland 20817-4874, USA. [12]Laboratory of Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA. [13]Manitoba Institute of Cell Biology, 675 McDermot Avenue, Room ON5029, Winnipeg, Manitoba R3E 0V9, Canada. [14]National Institute of Mental Health, 6001 Executive Blvd. - Rm 8235- MSC 9669, Bethesda, Maryland 20892-9669, USA. [15]Skirball Institute of Biomolecular Medicine, 540 First Avenue, 4th Floor, New York, New York 10016, USA. [16]Department of Genetics, University of North Carolina, CB 7248, 7007 Thurston Bowles Bldg, Chapel Hill, North Carolina 27599, USA. [17]School of Veterinary Medicine, University of California, One Shields Avenue, Davis, California 95616, USA. [18]Department of Genetics, Room 4109D Neurosciences Research Building, University of North Carolina, CB 7264, 103 Mason Farm Road, Chapel Hill, North Carolina 27599, USA. [19]Deltagen, 740 Bay Road, Redwood City, California 94063-2469, USA. [20]Samuel Lumenfeld Research Institute, University of Toronto, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. [21]National Institute on Drug Abuse, 6001 Executive Blvd, Rm 4274, Bethesda, Maryland 20892, USA. [22]GlaxoSmithKline, 5 Moore Drive, Durham, North Carolina 27709, USA. [23]Lexicon Genetics, 8800 Technology Forest Place, The Woodlands, Texas 77381-1160, USA. [24]Department of Molecular Biology, Massachusetts General Hospital, Wellman 911, 55 Fruit Street, Boston, Massachusetts 02114, USA. [25]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [26]The University of Tennessee-ORNL Graduate School of Genome Science and Technology, PO Box 2008, MS6164, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6164, USA. [27]Division of Basic Sciences, A2-025, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, P.O. Box 19024, Seattle, Washington 98109-1024, USA. [28]Cold Spring Harbor Laboratory, 1 Bungtown Road, PO Box 100, Cold Spring Harbor, New York 11724, USA. [29]Bioz, University of Technology, Dresden, c/o MPI-CBG, Pfotenhauerstr 108, 1307 Dresden, Germany. [30]Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA. [31]National Cancer Institute, National Institutes of Health, 31 Center Drive, Room 3A11, Bethesda, Maryland 20892-2440, USA. [32]Molecular Biology and Virology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037-1099, USA. [33]Merck Research Laboratories, PO Box 4, WP26-265, 770 Sumneytown Pike, West Point, Pennsylvania 19486, USA. [34]Laboratory for Molecular Hematology, University of Frankfurt Medical School, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany. [35]Banbury Center, Cold Spring Harbor Laboratory, PO Box 534, Cold Spring Harbor, New York 11724-0534, USA. [36]The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609, USA. [37]Institute of Developmental Genetics, GSF Research Center, Max-Planck-Institute of Psychiatry, Ingolstaeder Landstr. 1, 85764 Munich/Neuherberg, Germany. [38]Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, New York 10591, USA. [39]Gladstone Foundation for Cardiovascular Disease, University of California, San Francisco, California, USA. [40]Lexicon Genetics, 8800 Technology Forest Place, The Woodlands, Texas 77381-1160, USA. Correspondence should be addressed to C.P.A. (austinc@mail.nih.gov).

EXHIBIT C

Alberts, Molecular Biology of the Cell, p.543 (4th Ed. Garland Science (2002))

MOLECULAR BIOLOGY OF THE CELL

fourth edition

Garland
Vice President: Denise Schanck
Managing Editor: Sarah Gibbs
Senior Editorial Assistant: Kirsten Jenner
Managing Production Editor: Emma Hunt
Proofreader and Layout: Emma Hunt
Production Assistant: Angela Bennett
Text Editors: Marjorie Singer Anderson and Betsy Dilernia
Copy Editor: Bruce Goatly
Word Processors: Fran Dependahl, Misty Landers and Carol Winter
Designer: Blink Studio, London
Illustrator: Nigel Orme
Indexer: Janine Ross and Sherry Granum
Manufacturing: Nigel Eyre and Marion Morrow

Cell Biology Interactive
Artistic and Scientific Direction: Peter Walter
Narrated by: Julie Theriot
Production, Design, and Development: Mike Morales

Bruce Alberts received his Ph.D. from Harvard University and is President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. **Alexander Johnson** received his Ph.D. from Harvard University and is a Professor of Microbiology and Immunology at the University of California, San Francisco. **Julian Lewis** received his D.Phil. from the University of Oxford and is a Principal Scientist at the Imperial Cancer Research Fund, London. **Martin Raff** received his M.D. from McGill University and is at the Medical Research Council Laboratory for Molecular Cell Biology and Cell Biology Unit and in the Biology Department at University College London. **Keith Roberts** received his Ph.D. from the University of Cambridge and is Associate Research Director at the John Innes Centre, Norwich. **Peter Walter** received his Ph.D. from The Rockefeller University in New York and is Professor and Chairman of the Department of Biochemistry and Biophysics at the University of California, San Francisco, and an Investigator of the Howard Hughes Medical Institute.

Library of Congress Cataloging-In-Publicaton Data
Molecular biology of the cell / Bruce Alberts ... [et al.].-- 4th ed.
p. cm
Includes bibliographical references and index.
ISBN 0-8153-3218-1 (hardbound) -- ISBN 0-8153-4072-9 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology.]
QH581.2 .M64 2002
571.6--dc21

2001054471 CIP

Published by Garland Science, a member of the Taylor & Francis Group, 29 West 35th Street, New York, NY 10001-2299

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Front cover Human Genome: Reprinted by permission from *Nature*, International Human Genome Sequencing Consortium, 409:860–921, 2001 © Macmillan Magazines Ltd. Adapted from an image by Francis Collins, NHGRI; Jim Kent, UCSC; Ewan Birney, EBI; and Darryl Leja, NHGRI; showing a portion of Chromosome 1 from the initial sequencing of the human genome.

Back cover In 1967, the British artist Peter Blake created a design classic. Nearly 35 years later Nigel Orme (illustrator), Richard Denyer (photographer), and the authors have together produced an affectionate tribute to Mr Blake's image. With its gallery of icons and influences, its assembly created almost as much complexity, intrigue and mystery as the original. *Drosophila, Arabidopsis*, Dolly and the assembled company tempt you to dip inside where, as in the original, "a splendid time is guaranteed for all." (Gunter Blobel, courtesy of The Rockefeller University; Marie Curie, Keystone Press Agency Inc; Darwin bust, by permission of the President and Council of the Royal Society; Rosalind Franklin, courtesy of Cold Spring Harbor Laboratory Archives; Dorothy Hodgkin, © The Nobel Foundation, 1964; James Joyce, etching by Peter Blake; Robert Johnson, photo booth self-portrait early 1930s, © 1986 Delta Haze Corporation all rights reserved, used by permission; Albert L. Lehninger, (unidentified photographer) courtesy of The Alan Mason Chesney Medical Archives of The Johns Hopkins Medical Institutions; Linus Pauling, from Ava Helen and Linus Pauling Papers, Special Collections, Oregon State University; Nicholas Poussin, courtesy of ArtToday.com; Barbara McClintock, © David Micklos, 1983; Andrei Sakharov, courtesy of Elena Bonner; Frederick Sanger, © The Nobel Foundation, 1958.)

the target gene in a particular place or at a particular time. The most common of these recombination systems called **Cre/lox**, is widely used to engineer gene replacements in mice and in plants (see Figure 5–82). In this case the target gene in ES cells is replaced by a fully functional version of the gene that is flanked by a pair of the short DNA sequences, called lox sites, that are recognized by the Cre recombinase protein. The transgenic mice that result are phenotypically normal. They are then mated with transgenic mice that express the Cre recombinase gene under the control of an inducible promoter. In the specific cells or tissues in which Cre is switched on, it catalyzes recombination between the lox sequences—excising a target gene and eliminating its activity. Similar recombination systems are used to generate conditional mutants in *Drosophila* (see Figure 21–48).

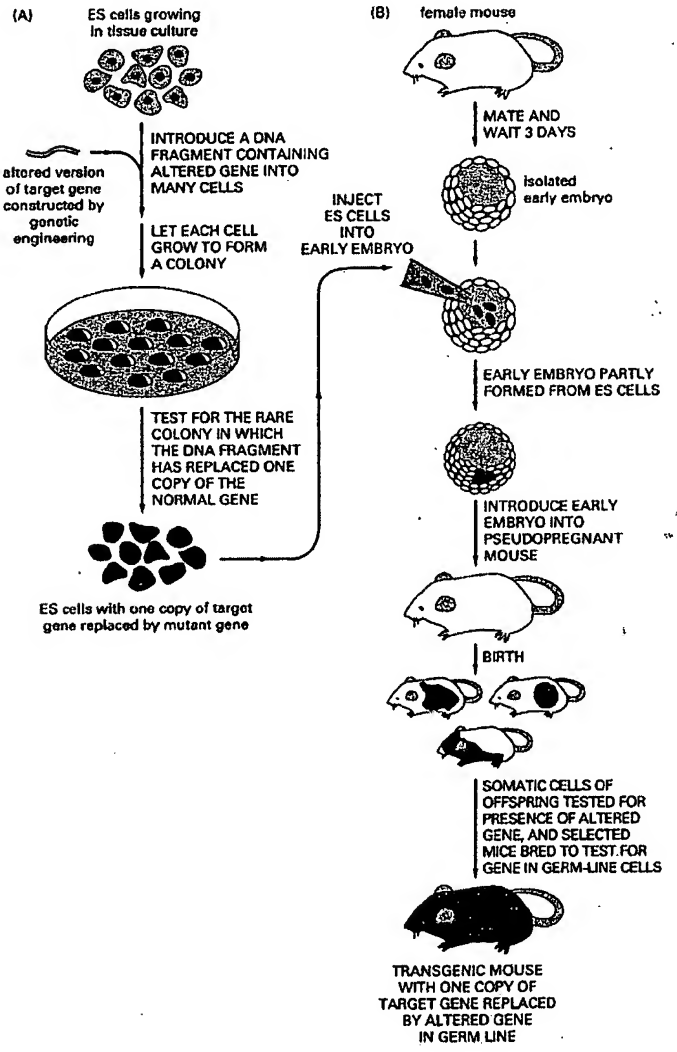


Figure 8–70 Summary of the procedures used for making gene replacements in mice. In the first step (A), an altered version of the gene is introduced into cultured ES (embryonic stem) cells. Only a few rare ES cells will have their corresponding normal genes replaced by the altered gene through a homologous recombination event. Although the procedure is often laborious these rare cells can be identified and cultured to produce many descendants, each of which carries an altered gene in place of one of its two normal corresponding genes. In the next step of the procedure (B), these altered ES cells are injected into a very early mouse embryo; the cells are incorporated into the growing embryo, and a mouse produced by such an embryo will contain some somatic cells (indicated by orange) that carry the altered gene. Some of these mice will also contain germ-line cells that contain the altered gene. When bred with a normal mouse, some of the progeny of these mice will contain the altered gene in all of their cells. If two such mice are in turn bred (not shown), some of the progeny will contain two altered genes (one on each chromosome) in all of their cells.

If the original gene alteration completely inactivates the function of the gene, these mice are known as knockout mice. When such mice are missing genes that function during development, they often die with specific defects long before they reach adulthood. These defects are carefully analyzed to help decipher the normal function of the missing gene.

Transgenic Plants Are Important for Both Cell Biology and Agriculture

When a plant is damaged, it can often repair itself by a process in which mature differentiated cells "dedifferentiate," proliferate, and then redifferentiate into other cell types. In some circumstances the dedifferentiated cells can even form an apical meristem, which can then give rise to an entire new plant, including gametes. This remarkable plasticity of plant cells can be exploited to generate transgenic plants from cells growing in culture.

When a piece of plant tissue is cultured in a sterile medium containing nutrients and appropriate growth regulators, many of the cells are stimulated to proliferate indefinitely in a disorganized manner, producing a mass of relatively undifferentiated cells called a callus. If the nutrients and growth regulators are carefully manipulated, one can induce the formation of a shoot and then root apical meristems within the callus, and, in many species, a whole new plant can be regenerated.

Callus cultures can also be mechanically dissociated into single cells, which will grow and divide as a suspension culture. In several plants—including tobacco, petunia, carrot, potato, and *Arabidopsis*—a single cell from such a suspension culture can be grown into a small clump (a clone) from which a whole plant can be regenerated. Such a cell, which has the ability to give rise to all parts of the organism, is considered **totipotent**. Just as mutant mice can be derived by genetic manipulation of embryonic stem cells in culture, so transgenic plants can be created from single totipotent plant cells transfected with DNA in culture (Figure 8–72).

The ability to produce transgenic plants has greatly accelerated progress in many areas of plant cell biology. It has had an important role, for example, in isolating receptors for growth regulators and in analyzing the mechanisms of morphogenesis and of gene expression in plants. It has also opened up many new possibilities in agriculture that could benefit both the farmer and the consumer. It has made it possible, for example, to modify the lipid, starch, and protein storage reserved in seeds, to impart pest and virus resistance to plants, and to create modified plants that tolerate extreme habitats such as salt marshes or water-stressed soil.

Many of the major advances in understanding animal development have come from studies on the fruit fly *Drosophila* and the nematode worm *Caenorhabditis elegans*, which are amenable to extensive genetic analysis as well as to experimental manipulation. Progress in plant developmental biology has, in the past, been relatively slow by comparison. Many of the plants that have proved most amenable to genetic analysis—such as maize and tomato—have long life cycles and very large genomes, making both classical and molecular genetic analysis time-consuming. Increasing attention is consequently being paid to a fast-growing small weed, the common wall cress (*Arabidopsis thaliana*), which has several major advantages as a "model plant" (see Figures 1–46 and 21–107). The relatively small *Arabidopsis* genome was the first plant genome to be completely sequenced.

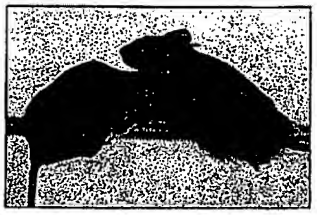


Figure 8–71 Mouse with an engineered defect in fibroblast growth factor 5 (FGF5). FGF5 is a negative regulator of hair formation. In a mouse lacking FGF5 (*right*), the hair is long compared with its heterozygous littermate (*left*). Transgenic mice with phenotypes that mimic aspects of a variety of human disorders, including Alzheimer's disease, atherosclerosis, diabetes, cystic fibrosis, and some type of cancers, have been generated. Their study may lead to the development of more effective treatments. (Courtesy of Gail Martin, from J.M. Hebert et al., *Cell* 78:1017–1025, 1994. © Elsevier.)

Large Collections of Tagged Knockouts Provide a Tool for Examining the Function of Every Gene in an Organism

Extensive collaborative efforts are underway to generate comprehensive libraries of mutations in several model organisms, including *S. cerevisiae, C. elegans, Drosophila, Arabidopsis*, and the mouse. The ultimate aim in each case is to produce a collection of mutant strains in which every gene in the organism has either been systematically deleted, or altered such that it can be conditionally disrupted. Collections of this type will provide an invaluable tool for investigating gene function on a genomic scale. In some cases, each of the individual mutants within the collection will sport a distinct molecular tag—a unique DNA sequence designed to make identification of the altered gene rapid and routine. In *S. cerevisiae*, the task of generating a set of 6000 mutants, each missing

EXHIBIT D

Genes VII, p.508 (Lewin, Oxford University Press (2000))

Genes VII

Benjamin Lewin

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Athens Auckland Bangkok Bogotá Buenos Aires Calcutta
Cape Town Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul
Karachi Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi Paris São Paulo Singapore Taipei Tokyo Toronto Warsaw
with associated companies in Berlin Ibadan

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2000

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
(Data applied for)

ISBN 0-19-879276-X (Hbk)

Typeset by J&L Composition Ltd, Filey, North Yorkshire
Printed in The United States of America

which exogenous DNA is introduced from a bacterium into a host cell. The mechanism resembles that of bacterial conjugation. Expression of the bacterial DNA in its new host changes the phenotype of the cell. In the example of the bacterium *Agrobacterium tumefaciens*, the result is to induce tumor formation by an infected plant cell.

Alterations in the relative proportions of components of the genome during somatic development occur to allow insect larvae to increase the number of copies of certain genes. And the occasional amplification of genes in cultured mammalian cells is indicated by our ability to select variant cells with an increased copy number of some genes. Initiated within the genome, the amplification event can create additional copies of a gene that survive in either intrachromosomal or extrachromosomal form.

When extraneous DNA is introduced into eukaryotic cells, it may give rise to extrachromosomal forms or may be integrated into the genome. The relationship between the extrachromosomal and genomic forms is irregular, depending on chance and to some degree unpredictable events, rather than resembling the regular interchange between free and integrated forms of bacterial plasmids.

Yet, however accomplished, the process may lead to stable change in the genome; following its injection into animal eggs, DNA may even be incorporated into the genome and inherited thereafter as a normal component, sometimes continuing to function. Injected DNA may enter the germline as well as the soma, creating a transgenic animal. The ability to introduce specific genes that function in an appropriate manner could become a major medical technique for curing genetic diseases.

The converse of the introduction of new genes is the ability to disrupt specific endogenous genes. Additional DNA can be introduced within a gene to prevent its expression and to generate a null allele. Breeding from an animal with a null allele can generate a homozygous "knockout", which has no active copy of the gene. This is a powerful method to investigate directly the importance and function of a gene.

Considerable manipulation of DNA sequences therefore is achieved both in authentic situations and by experimental fiat. We are only just beginning to work out the mechanisms that permit the cell to respond to selective pressure by changing its bank of sequences or that allow it to accommodate the intrusion of additional sequences.

The mating pathway is triggered by signal transduction

THE yeast *S. cerevisiae* can propagate happily in either the haploid or diploid condition. Conversion between these states takes place by mating (fusion of haploid spores to give a diploid) and by sporulation (meiosis of diploids to give haploid spores). The ability to engage in these activities is determined by the mating type of the strain.

The properties of the two mating types are summarized in Figure 17.1. We may view them as resting on the teleological proposition that there is no point in mating unless the haploids are of different genetic types; and sporulation is productive only when the diploid is heterozygous and thus can generate recombinants.

The mating type of a (haploid) cell is determined by the genetic information present at the *MAT* locus. Cells that carry the *MATa* allele at this locus are type a; likewise, cells that carry the *MATα* allele are type α. Cells of opposite type can mate; cells of the same type cannot.

Recognition of cells of opposite mating type is

Figure 17.1 Mating type controls several activities.

	MATa	*MATα*	*MATa/MATα*
Cell type	a	α	a/α
Mating	yes	yes	no
Sporulation	no	no	yes
Pheromone	a factor	α factor	none
Receptor	binds α factor	binds a factor	none

EXHIBIT E

Joyner, Gene Targeting: A Practical Approach, preface, Oxford University Press (2000)

Gene Targeting

A Practical Approach

Second Edition

Edited by

ALEXANDRA L. JOYNER

Howard Hughes Medical Institute and Skirball Institute of Biomolecular Medicine, New York University School of Medicine

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford and furthers the University's aim of excellence in research, scholarship, and education by publishing worldwide in
Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi São Paulo Shanghai Taipei Tokyo Toronto
Oxford is a registered trade mark of Oxford University Press
Published in the United States by Oxford University Press Inc., New York

First published 2000
Reprinted 2001, 2003

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data

Gene targeting : a practical approach / edited by Alexandra L. Joyner.
—2nd ed.
Includes bibliographical references and index.
1. Gene targeting Laboratory manuals. I. Joyner, Alexandra L.
QH442.3.G456 1999 660.6'5—dc21 99–36660

ISBN 0-19-963792-X (Pbk)

Printed in Great Britain by Information Press, Ltd, Eynsham, Oxon.

Preface

Over the past ten years it has become possible to make essentially any mutation in the germline of mice by utilizing recombination and embryonic stem (ES) cells. Homologous recombination when applied to altering specific endogenous genes, referred to as gene targeting, provides the highest level of control over producing mutations in cloned genes. When this is combined with site specific recombination, a wide range of mutations can be produced. ES cell lines are remarkable since after being established from a blastocyst, they can be cultured and manipulated relatively easily in vitro and still maintain their ability to step back into a normal developmental program when returned to a pre-implantation embryo. With the exponential increase in the number of genes identified by various genome projects and genetic screens, it has become imperative that efficient methods be developed for determining gene function. Gene targeting in ES cells offers a powerful approach to study gene function in a mammalian organism. Gene trap approaches in ES cells, in particular when they are combined with sophisticated prescreens, offer not only a route to gene discovery, but also to gain information on gene sequence, expression and mutant phenotype.

The basic technology necessary for making designer mutations in mice has become widespread and researchers who have traditionally used cell biology or molecular experiments are adding gene targeting techniques to their repertoire of experimtal approaches. A second edition of this book was written for two main reasons. The first was to update previously described techniques and to add new techniques that have greatly expanded the types of mutations that can be made using recombination in ES cells. A chapter in this new edition describes the design and use of site specific recombination for gene targeting approaches and production of conditional mutations. The second reason for the new book was to provide a more in depth discussion of the experimental design considerations that are critical to a successful gene targeting study and to add approaches for analyzing mutant phenotypes, the most interesting part of an experiment. Gene targeting experiments should be designed to go far beyond just making a mutant mouse. The success of a gene targeting experiment no longer lies in the making of the mutation, but depends on the imaginative and insightful analysis of the mutant phenotypes that the mutation provides. A chapter in this edition describes the use of classical genetics in combination with gene targeting to get the most out of a genetic approach to a biological question.

The nature of in vivo gene targeting studies of gene function are such that critical design decisions must be made at every step in the experiment, and each decision can have a major impact on the value of the information obtained. From the start, the type of mutation to be made must be considered

carefully. Whereas 10 years ago most mutations were designed to create null mutations and were therefore relatively simple to design, at present, a null mutation is only one of a long list of mutations that can be made, each providing different insight into the function of a gene. Point mutations, large deletions, gene exchanges (knock-ins) and conditional mutations are but a few of the choices one faces at the start of a gene targeting experiment. The next choice is the source of DNA for the targeting experiment and ES cell line to be used for the manipulations. Once the mutant ES cell clone has been obtained, there are then a number of alternative approaches that can be used to make ES cell chimeras that depend on the ES cell line which was used. Finally, and most importantly, is the analysis of any phenotype that arises. This second addition discusses techniques used to analyze mutant mice, ranging from standard descriptive evaluation, to a chimera analysis or complicated breeding experiments that utilize double mutants. If mice are simply considered as a 'bag of cells' or an in vivo source of selected cell types, then the tremendous resource which mice offer as a model organism is not being realized. The life of a mouse represents a continuum of dynamic processes, including pattern formation, organ development, learning, homeostasis and disease. By making genetic alterations in mice using gene targeting and ES cells, the effects of a given change can be studied in the context of the whole organism.

My goal in editing this book was to provide a manual that could take a newcomer to the exciting field of gene targeting and mutant analysis in mice from a cloned gene to a basic understanding of the genetic approaches available using ES cells, and how each technique can be used to design a particular in vivo test of gene function. The book should also provide a valuable bench side resource for anyone carrying out gene targeting or gene trap experiments, a chimera analysis or classical genetic approaches. I would once again like to extend many thanks and my deepest appreciation to all the authors for their great efforts in including detailed protocols and lucid discussions of the various approaches presented. I would also like to thank my family for their strong support and laboratory members past and present for helping to make gene targeting a reality. Finally, since many of the techniques use mice, the experiments should be carried out in accordance with local regulations.

New York, NY A.L.J.

1

Gene targeting, principles, and practice in mammalian cells

PAUL HASTY, ALEJANDRO ABUIN, and ALLAN BRADLEY

1. Introduction

When a fragment of genomic DNA is introduced into a mammalian cell it can locate and recombine with the endogenous homologous sequences. This type of homologous recombination, known as gene targeting, is the subject of this chapter. Gene targeting has been widely used, particularly in mouse embryonic stem (ES) cells, to make a variety of mutations in many different loci so that the phenotypic consequences of specific genetic modifications can be assessed in the organism.

The first experimental evidence for the occurrence of gene targeting in mammalian cells was made using a fibroblast cell line with a selectable artificial locus by Lin *et al.* (1), and was subsequently demonstrated to occur at the endogenous β-globin gene by Smithies *et al.* in erythroleukaemia cells (2). In general, the frequencies of gene targeting in mammalian cells are relatively low compared to yeast cells and this is probably related to, at least in part, a competing pathway: efficient integration of the transfected DNA into a random chromosomal site. The relative ratio of targeted to random integration events will determine the ease with which targeted clones are identified in a gene targeting experiment. This chapter details aspects of vector design which can determine the efficiency of recombination, the type of mutation that may be generated in the target locus, as well as the selection and screening strategies which can be used to identify clones of ES cells with the desired targeted modification. Since the most common experimental strategy is to ablate the function of a target gene (*null allele*) by introducing a selectable marker gene, we initially describe the vectors and the selection schemes which are helpful in the identification of recombinant clones (Sections 2–5). In Section 6, we describe the vectors and additional considerations for generating subtle mutations in a target locus devoid of any exogenous sequences. Finally, Section 7 is dedicated to the use of gene targeting as a method to express exogenous genes from specific endogenous regulatory elements *in vivo*, also known as 'knock-in' strategies.

enrich populations of transfected cells for targeted integration events (Section 4.2.1).

2.1 Design considerations of a replacement vector

The principal consideration in the design of a replacement vector, is the type of mutation generated. Secondary (yet still important) considerations relate to the selection scheme and screening techniques required to isolate the recombinant clones. The recombinant alleles generated by replacement vectors typically have a selection cassette inserted into a coding exon or replacing part of the locus. It is important to consider that, exon interruptions and small deletions will not necessarily ablate the function of the target gene to generate a null allele. Consequently, it is necessary to confirm that the allele which has been generated is null by RNA and/or protein analysis and in many cases transcripts and truncated proteins from such a mutant allele can be detected. Considering that products from the mutated locus may have some function (normal or abnormal) it is important to design a replacement vector so that the targeted allele is null, particularly in the absence of a good assay for the gene product. Disruption or deletion of the coding sequence by the positive selection marker will in most instances ablate a gene's function. However in some situations a truncated protein may be generated which retains some biological activity, thus some knowledge of mutations in a related gene in another organism can be helpful in the determination of the possible function of a targeted allele. Null alleles are more likely to occur by deleting or recombining a selection cassette into more 5' exons rather than exons that encode the C-terminus of the protein, since under these circumstances minimal portions of the wild-type polypeptide would be made.

There are several considerations to take into account when a positive selection marker is to be inserted into an exon. One critical consideration is that since the length of an exon can influence RNA splicing (3), an artificially large exon caused by the insertion of a selectable marker may not be recognized by the splicing machinery and could be skipped. Thus, transcripts initiated from the endogenous promoter may delete the mutated exon from the mRNA species or even additional exons. If a skipped exon is a coding exon whose nucleotide length is not a multiple of three (codon) the net result will be both a deletion and a frame-shift mutation of the gene, which will often generate a null allele. However, if the disrupted coding exon has a nucleotide length which is a multiple of three, if spliced out, this would result in a protein with a small in-frame deletion which may retain partial or complete function. The same concept applies to gene targeting vectors in which exons are being deleted and replaced by the selectable marker. Deletion of an exon or group of exons with a unit number of codons may also result in a functional protein product with an in-frame deletion. For most purposes it is advisable to delete portions or all of the target gene so that the genetic

EXHIBIT F

Matise et al., *Production of Targeted Embryonic Stem Cell Clones*, p.101
in
Joyner, *Gene Targeting: A Practical Approach*, Oxford University Press 2000)

3

Production of targeted embryonic stem cell clones

MICHAEL P. MATISE, WOJTEK AUERBACH and ALEXANDRA L. JOYNER

1. Introduction

The discovery that cloned DNA introduced into tissue culture cells can undergo homologous recombination at specific chromosomal loci has revolutionized our ability to study gene function in cell culture and *in vivo*. In theory, this technique, termed gene targeting, allows one to generate any type of mutation in any cloned gene. The kinds of mutations that can be created include null mutations, point mutations, deletions of specific functional domains, exchanges of functional domains from related genes, and gain-of-function mutations in which exogenous cDNA sequences are inserted adjacent to endogenous regulatory sequences. In principle, such specific genetic alterations can be made in any cell line growing in culture. However, not all cell types can be maintained in culture under the conditions necessary for transfection and selection. Over ten years ago, pluripotent embryonic stem (ES) cells derived from the inner cell mass (ICM) of mouse blastocyst stage embryos were isolated and conditions defined for their propogation and maintenance in culture (1, 2). ES cells resemble ICM cells in many respects, including their ability to contribute to all embryonic tissues in chimeric mice. Using stringent culture conditions, the embryonic developmental potential of ES cells can be maintained following genetic manipulations and after many passages *in vitro*. Furthermore, permanent mouse lines carrying genetic alterations introduced into ES cells can be obtained by transmitting the mutation through the germline by generating ES cell chimeras (described in Chapters 4 and 5). Thus, applying gene targeting technology to ES cells in culture affords researchers the opportunity to modify endogenous genes and study their function *in vivo*. In initial studies, one of the main challenges of gene targeting was to distinguish the rare homologous recombination events from more commonly occurring random integrations (discussed in Chapter 1). However, advances in cell culture and in selection schemes, in vector construction using isogenic DNA, and in the application of rapid screening procedures have made it possible to identify homologous recombination events efficiently.

EXHIBIT G

Crawley, *What's Wrong With My Mouse Behavioral Phenotyping of Transgenic and Knockout Mice,* p.1, rear cover, Wiley-Liss 2000)

What's Wrong With My Mouse?

Behavioral Phenotyping of Transgenic and Knockout Mice

Jacqueline N. Crawley, Ph.D.

WILEY-LISS

A John Wiley & Sons, Inc., Publication

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

This book was written by Dr. Jacqueline Crawley in her private capacity, outside of her professional position as Chief, Section on Behavioral Neuropharmacology, National Institute of Mental Health, Bethesda, Maryland, USA. The views expressed in this book do not necessarily represent the views of the National Institutes of Health or of the United States government.

This book is printed on acid-free paper.

Published simultaneously in Canada.

For ordering and customer service, call 1-800-CALL-WILEY.

Library of Congress Cataloging-in-Publication Data:
Crawley, Jacqueline N.
What's wrong with my mouse? : behavioral phenotyping of transgenic and knockout mice / Jacqueline N. Crawley.
p. cm.
Includes bibliographical references.
ISBN 0-471-31639-3 (cloth : alk. paper)
1. Neurogenetics—Animal models. 2. Transgenic mice—Behavior.
I. Title.
QP356.22.C73 1999
573.8'48'0724—dc21 99-28955

Printed in the United States of America.
10 9 8 7 6 5 4

Preface

Targeted mutation of genes expressed in the nervous system is an exciting new research field that is forging a remarkable amalgam of molecular genetics and behavioral neuroscience. My laboratory in Bethesda has been the fortunate recipient of visits from many molecular geneticists over the past five years, who come to ask, "What's wrong with my mouse? Can you tell us what behaviors are abnormal in our null mutants? And how *do* you measure behavior, anyway?"

We have had some remarkable opportunities to collaborate with outstanding molecular geneticists in the National Institutes of Health Intramural Research Program and throughout the world on investigations of the behavioral effects of mutations in genes expressed in the mouse brain. Each of these collaborations has been a learning experience, increasing our understanding of the optimal experimental design for analyzing behavioral phenotypes of mutant mice. What are the best tests to address each specific hypothesis? Which methods work best for mice? Which rat tasks can be adapted for mice? What are the correct controls? What are the hidden pitfalls, lurking artifacts, false positives, and false negatives? Which statistical tests are most sensitive for detection of the genotype effect? What is the minimum number of animals necessary for each genotype, gender, and age? Our laboratory and many others are gradually working out the best methods for behavioral phenotyping of transgenic and knockout mice.

In the same conversations, molecular geneticists frequently asked me to recommend a book they could consult to learn more about behavioral tests for mice. Apparently the scientific book publishers are receiving similar queries. Ann Boyle and Robert Harington at John Wiley & Sons, convinced of a real need for such a book, sweet-talked me into filling the void. *What's Wrong With My Mouse?* is written for these pioneering molecular geneticists, and for the talented students who will be the next driving force in moving the field forward.

On a personal level, I would like to express deep appreciation to all of my behavioral neuroscientist colleagues around the world for their outstanding work, past, present, and future. Your contributions to the excellence and abundance of mouse behavioral tests provide

the foundation for the rapidly expanding scientific discoveries forthcoming from behavioral phenotyping studies of transgenic and knockout mice. This book is a testament to your accomplishments.

JACQUELINE N. CRAWLEY, PH.D.
Chevy Chase, Maryland

1

Designer Mice

The disease is inherited. Family pedigrees indicate an autosomal dominant gene. Linkage analyses reveal one strongly associated chromosomal locus. Mapping identifies the gene. The cDNA for the gene is sequenced. The anatomical distribution of the gene is primarily in the brain. The symptoms of the disease are primarily neuropsychiatric. There is no treatment for the disease. The disease is lethal.

Your mission, should you choose to accept it, is to develop a treatment for the disease. Replacement gene therapy is the best hope. But you don't know the gene product, you don't know its function, and you don't know if gene delivery would be therapeutic. Where do you start?

These days, you may choose to start with a targeted gene mutation, to generate a mutant mouse model of the hereditary disease. A DNA construct containing the mutated form of the responsible gene is developed. The construct is inserted into the mouse genome. A line of mice with the mutated gene is generated. Characteristics of the mutant mice are identified in comparison to normal controls. Salient characteristics relevant to the human disease are quantitated. These diseaselike traits are then used as test variables for evaluating the effectiveness of treatments. Putative treatments are administered to the mutant mice. A treatment that prevents or reverses the disease traits in the mutant mice is taken forward for further testing as a potential therapeutic treatment for the human genetic disease. Gene therapy, based on targeted gene replacement of the missing or incorrect gene in the human hereditary disease, is described in Chapter 12. In the future, medicine may shift emphasis from treating the symptoms to administering replacement genes that effectively and permanently cure the disease.

Targeted gene mutation in mice represents a new technology that is revolutionizing biomedical research. Transgenic mice have an *extra gene* added. An additional copy of a normal gene is inserted into the mouse genome to study overexpression of the gene product. Or a new gene is added that is not normally present in the mouse genome. The new gene may be the aberrant form of a human gene linked to a disease. For example, the mutated form of the human *huntingtin* gene is added to the mouse genome to generate a mouse model of Huntington's disease. Knockout mice have a *gene deleted*. The null mutant homozygous

knockout mouse is deficient in both alleles of a gene; the heterozygote is deficient in one of its two alleles for the gene. The genotype is −/− for the null mutant, +/− for the heterozygote, and +/+ for the wildtype normal control. The phenotype is the set of observed characteristics resulting from the mutation. Phenotypes include biochemical, anatomical, physiological, and behavioral characteristics.

Targeted mutations of genes expressed in the brain are revealing the mechanisms underlying normal behavior and behavioral abnormalities. Mouse models of human neuropsychiatric diseases, such as Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, obesity, anorexia, depression, alcoholism, drug addiction, schizophrenia, and anxiety, are likely to be characterized by their behavioral phenotype.

This book is designed to introduce the novice to the rich literature of behavioral tests in mice and to show how to optimize the application of these tests for behavioral phenotyping of mutant mice. Based on our experiences, our laboratory is working toward a unified approach for the optimal conduct of behavioral phenotyping experiments in mutant mice. Recommendations are offered for a three-tiered sequence of behavioral tests, applicable to each behavioral domain relevant to genes expressed in the mammalian brain.

SCOPE

This book is designed as an overview of the mutant mouse technology and an introducion to the field of behavioral neuroscience, as it can be applied to behavioral phenotyping of transgenic and knockout mice. Molecular geneticists may browse through the chapters relevant to their gene, to get ideas for possible tests to try. Behavioral neuroscientists who have no experience with mutant mice may wish to read about the methods for developing a transgenic or knockout, the behavioral tests that have been effectively applied, and some of the successful experiments published in the genetics literature.

Chapters are organized around behavioral domains, including general health, neurological reflexes, developmental milestones, motor functions, sensory abilities, learning and memory, feeding, sexual and parental behaviors, social behaviors, and rodent paradigms relevant to fear, anxiety, depression, schizophrenia, reward, and drug addiction. Each chapter begins with a brief history of the early work in the field and the present hypotheses about mechanisms underlying the expression of the behavior. A list of general review articles and books is offered for each topic, encouraging the interested reader to gain more in-depth knowledge of the relevant literature.

Standard tests are then presented in detail. Highlighted are those tasks that have been extensively validated in mice. Demonstrations of genetic components of task performance are described, including experiments comparing inbred strains of mice (strain distributions), quantitative trait loci approaches (linkage analysis), and naturally occurring mutants (spontaneous mutations). Experimental design and specific behavioral tasks are presented as simply as possible. Extensive references are included for each behavioral test to obtain more complete methods from the primary experimental literature on the topic.

Illustrations are provided for the most frequently used behavioral tasks. Photographs of the equipment or diagrams of the task accompany the text. Samples of data are shown. The data presentation is designed to indicate the qualitative and quantitative results that can be expected when the task is properly conducted.

Each chapter includes the results of several representative experiments in which these tasks are successfully applied to characterize transgenic and knockout mice. Examples are

WHAT'S WRONG WITH MY MOUSE ?

Transgenic and knockout mutations provide an important means for understanding gene function, as well as for developing therapies for genetic diseases. This engaging and informative book discusses the many advances in the field of transgenic technology that have enabled researchers to bring about various changes in the mouse genome. Equal emphasis is given to both the principles of transgenic and knockout methods and their applications. A clear and concise format provides researchers with a comprehensive review of the behavioral paradigms appropriate for analyzing mouse phenotypes.

What's Wrong with My Mouse? explains the differences between transgenic knockout mice and their wild-type controls, while providing critical information about gene function and expression. This volume recognizes that newly identified genes can provide useful insights into brain functioning, including brain malfunctioning in disease states. Written by a world-renowned expert in the field, the material also covers:

- How to generate a transgenic or knockout mouse
- Motor functions (open field, holeboard, rotarod, balance, grip, circadian activity, etc.)
- Sensory abilities (olfaction, vision, hearing, taste, touch, nociception)
- Reproductive behavior, social behavior, and emotional behavior

Researchers in neuroscience, pharmacology, genetics, developmental biology, and cell biology will all find this book essential reading.

Cover Design: Paul DiNovo

WILEY-LISS

A JOHN WILEY & SONS, INC., PUBLICATION

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

ISBN 0-471-31639-3

9 780471 316398 90000

EXHIBIT H

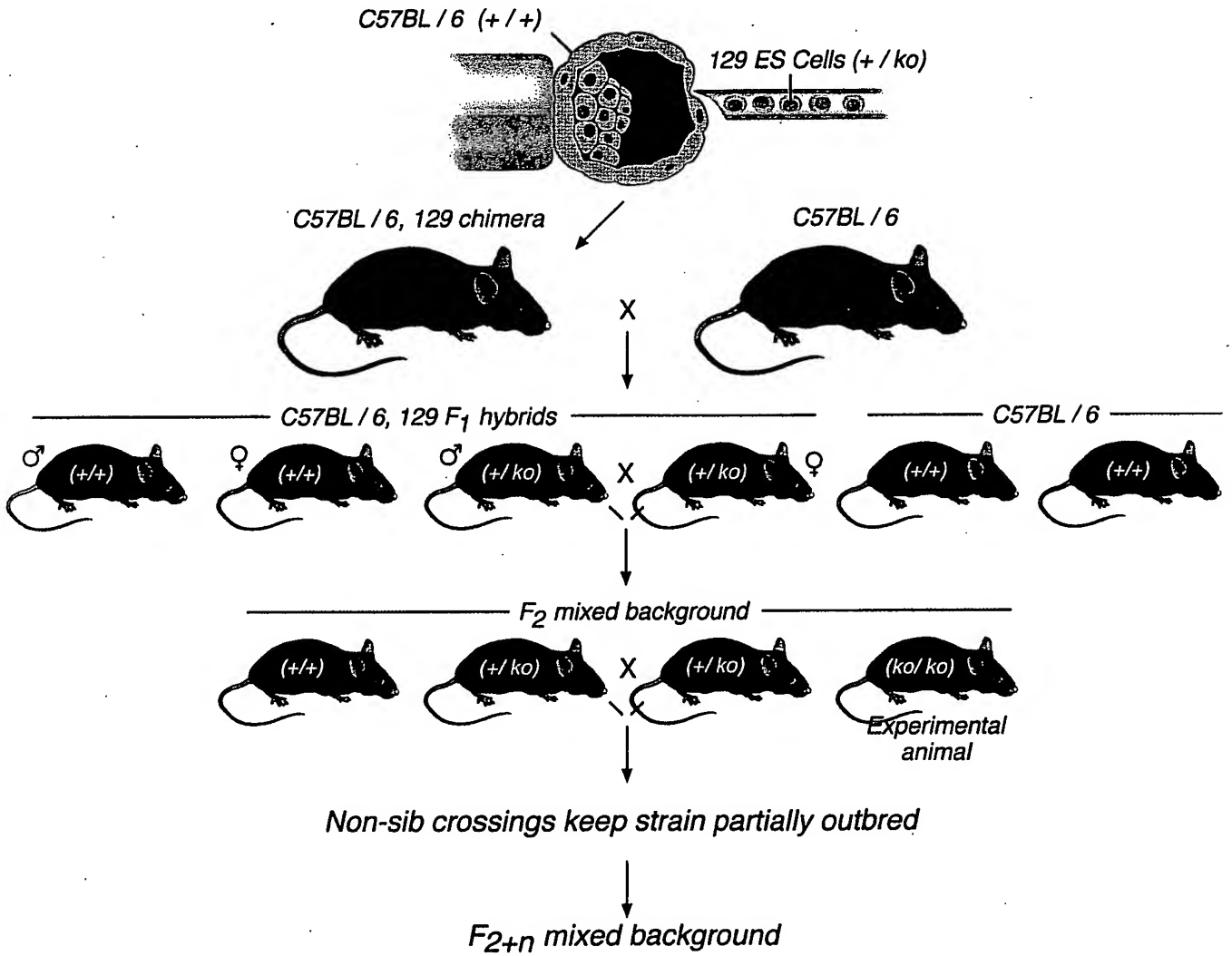
Doetschman, Laboratory Animal Science 49:137-143, 137 (1999)

Special Topic Overview:
Interpretation of Phenotype in Genetically Engineered Mice
Volume 49, Number 2

Laboratory Animal Science

International Journal of Comparative and Experimental Medicine

Knockout Mice on a Mixed Genetic Background



F_{2+n} (ko/ ko) experimental animals on mixed C57BL / 6, 129 background can provide a wide range of phenotypes of varying penetrance and expressivity, reflecting the human condition.

American Association for Laboratory Animal Science

Laboratory Animal Science
Copyright 1999
by the American Association for Laboratory Animal Science

Vol 49, No 2
April 1999

Special Topic Overview

Interpretation of Phenotype in Genetically Engineered Mice

Thomas Doetschman

Background and Purpose: **In mice, genetic engineering involves two general approaches—addition of an exogenous gene, resulting in transgenic mice, and use of knockout mice, which have a targeted mutation of an endogenous gene. The advantages of these approaches is that questions can be asked about the function of a particular gene in a living mammalian organism, taking into account interactions among cells, tissues, and organs under normal, disease, injury, and stress situations.**

Methods: **Review of the literature concentrating principally on knockout mice and questions of unexpected phenotypes, lack of phenotype, redundancy, and effect of genetic background on phenotype will be discussed.**

Conclusion: **There is little gene redundancy in mammals; knockout phenotypes exist even if none are immediately apparent; and investigating phenotypes in colonies of mixed genetic background may reveal not only more phenotypes, but also may lead to better understanding of the molecular or cellular mechanism underlying the phenotype and to discovery of modifier gene(s).**

One often hears the comment that genetically engineered mice, especially knockout mice, are not useful because they frequently do not yield the expected phenotype, or they don't seem to have any phenotype. These expectations are often based on years of work, and in some instances, thousands of publications of mostly in vitro studies. Examples of unexpected phenotypes, based largely on experience with transforming growth factor beta (*Tgfb*) and basic fibroblast growth factor (*Fgf2*) knockout and transgenic mice, will be presented to discuss possible reasons for unexpected knockout phenotypes. The conclusions will be that the knockout phenotypes do, in fact, provide accurate information concerning gene function, that we should let the unexpected phenotypes lead us to the specific cell, tissue, organ culture, and whole animal experiments that are relevant to the function of the genes in question, and that the absence of phenotype indicates that we have not discovered where or how to look for a phenotype.

Before entering into how one should interpret unexpected knockout phenotypes and how one should deal with lack of knockout phenotypes, it is necessary to give a brief introduction into how knockout mice are made. For detailed information, the following reviews are suggested (1–4). Transgenic technology has had a long history; thus, an introduction to that technology will not be given here. Rather, the following reviews are suggested (5, 6). At this juncture, it should be noted that, although transgenic vertebrates ranging from fish to bovids have been produced, knockout technology has to date been successful only in mice, even though embryonic stem (ES) cells have been produced from several other species, including hamster (7), rat (8), rabbit (9, 10), pig (11–13), bovine (14, 15), and zebrafish (16). Consequently, the entire discussion will be focused on mice.

Knockout mice are generated by the injection of genetically engineered or gene-targeted ES cells into a mouse blastocyst to generate a chimeric embryo, which in turn can pass on the engineered gene to its offspring. ES cell lines are established from the inner cell mass of a mouse blastocyst, so that when injected into blastocysts, the ES cells can incorporate into the inner cell mass of the recipient blastocysts, thereby chimerizing them. Subsequent to transfer of the chimeric blastocysts into uteri of pseudopregnant mice, chimeric mice are born. If the germline of a chimeric mouse is colonized by cells derived from the injected ES cells, the chimera is termed a "germline" chimera. Some of the offspring of the germline chimeras will then carry the engineered gene in their genomes. Gene targeting in ES cells uses the ES cells' DNA repair apparatus to bring about homologous recombination between an exogenous DNA fragment transfected into the ES cell and its homologous region in the genome. Homologous recombination usually results in replacement of the endogenous region with the exogenous fragment, thereby altering the endogenous gene in a prespecified manner. There are many variations on this procedure by which genes can be altered not only to ablate function, but also to make more subtle mutations (17–19). Such procedures can be used to introduce point mutations, remove specific splicing products, switch isoforms, and humanize genes. In addition, technology has recently been

Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, Ohio

developed to make conditional and inducible knockouts in which gene function is ablated either in a developmentally specified tissue (20–22) or in an inducible manner (23–26). These techniques, though exciting, will not be further discussed.

Extensive nonredundancy in the TGFβ family: Several thousand cell culture studies on the three mammalian transforming growth factor beta proteins (TGFβs 1, 2, and 3) have implicated these growth and differentiation factors in the function of nearly every cell type studied. Expression studies indicated unique and overlapping expression of the three TGFβs (27, 28). For example, overlapping protein localization was found in all gut epithelia, all layers of the skin, all three muscle types, kidney tubules, lung bronchi, cartilage, and bone (Table 1). Together with the fact that all three TGFβs signal through a common TGF type-II receptor (Figure 1), these data strongly suggest considerable redundancy in function. Consequently, it is surprising that, of the >30 phenotypes of the three *Tgfb* knockout mice that we have described (29–31), none appear to be overlapping (Table 2). These results indicate extensive nonredundancy between TGFβ ligands even though there is considerable overlap in expression. Of course, these results do not rule out the possibility of some redundancy in some tissues. Combination of the ligand knockouts would uncover such situations, and it is likely that a few will exist, but 30 nonoverlapping phenotypes for three ligands strongly suggests that a vast number of their functions are not redundant.

There are several possible explanations for how there can be so much overlap in ligand expression and yet so much specific ligand function. First, TGFβs are secreted as latent peptides and must be activated before they can bind receptors (32–35). The mechanism by which this extracellular processing occurs is not well understood and may be different for each TGFβ. Hence, ligand processing presumably determines some functional specificity for the three TGFβs. Second, there is a third type of TGFβ receptor, TGFβR3, that can interact with ligand and receptor types I and II before cytoplasmic signaling can occur, though involvement of TGFβR3 is not essential for signaling (36–38). Association with type III receptors is thought to enhance some TGFβR1 and 2/ligand interactions. Upon ligand binding, the serine/threonine receptor TGFβR2 then associates with and phosphorylates the transmembrane serine/threonine receptor TGFβR1, which in turn initiates a phosphorylation-mediated signaling cascade. Hence, combinatorial receptor/ligand interactions will also determine functional specificity. Third, signaling from TGFβR1 can occur through two cytoplasmic signaling proteins called SMAD2 and 3 (39, 40) and, perhaps, through a third called SMAD5 (41). In addition, SMAD6 and 7 can also interact with the other SMADs to inhibit signaling (42–44). Hence, differential SMAD protein interactions with transcriptional machinery will probably also determine functional specificity for the three TGFβ ligands. Finally, there may be several non-transcriptional signaling pathways for TGFβs. For example, we have found that TGFβ1-deficient platelets from *Tgfb1* knockout mice have impaired platelet aggregation that can be restored by incubating isolated platelets with recombinant TGFβ1 (unpublished observations). Because platelets do not have a nucleus, there must exist a signaling pathway that is nontranscriptional. In summary, given the complexities of ligand processing, receptor interactions, and signaling pathways, it becomes clear why redundancy in TGF1, 2, and 3 function has not been detected at the whole animal level, even though there is considerable overlap in expression of *Tgfb* gene family members. Consequently, if other gene families function with similar complexity, it is likely that, in the final analysis, little functional redundancy will be found within gene families.

Two striking examples of apparent functional redundancy are worth considering. The first involves myogenic genes, and the second involves retinoic acid receptors. Contrary to early interpretations, redundancy does not now appear to be

Table 1. Protein expression of transforming growth factor beta (TGFβ) 1, 2, and 3

Tissue/cell type	TGFβ1	TGFβ2	TGFβ3
Cartilage			
Perichondrium	+++	+	++
Chondrocytes	+	++	++
Bone			
Periosteum	++	-	+
Osteocytes		++	++ ++
Tooth			
Ameloblasts	+	-	+
Odontoblasts	-	++	-
Pulp	+	+++	+
Muscle			
Smooth	+	+	++
Cardiac	+	+	+++
Skeletal	+	++	
Lung			
Bronchi	++	++	++
Alveoli	-	-	-
Blood vessels			
Endothelium	-	-	++
Smooth muscle	+	+	+++
Kidney			
Tubules	++	++	++
Basement membrane	-	+++	-
Adrenal			
Cortex	+++	+++	-
Medulla	-	-	-
Gut			
Esophageal epithelium	+++	+	+
Gastric epithelium	+++	+	+
Intestinal epithelium	++	+	+
Basement membrane	-	+++	-
Muscularis	+	+	++
Liver			
Capsule	-	-	++
Parenchyma	-	-	-
Megakaryocytes	+	-	++
Eye			
Lens epithelium	-	-	-
Lens fibers	+++	+	+
Ear			
Cochlear epithelium	-	+	+++
Basement membrane	-	+++	-
CNS			
Meninges	+	+++	+
Glia	-	++	++
Choroid plexus	-	-	++
Skin			
Periderm	++	+	++
Epidermis	+++	+++	+++
Dermis	+	+++	+
Hair follicles	++	++	+

The polyclonal antibodies used were specific for residues 4–19 of TGFβ1 and 2 and residues 9–20 of TGFβ3. The avidin-biotin system was used for staining. Data obtained from immunohistochemical study of Pelton et al. (28). Reproduced from *The Journal of Cell Biology*, 1991,115:1091–1105, by copyright permission of The Rockefeller University Press.

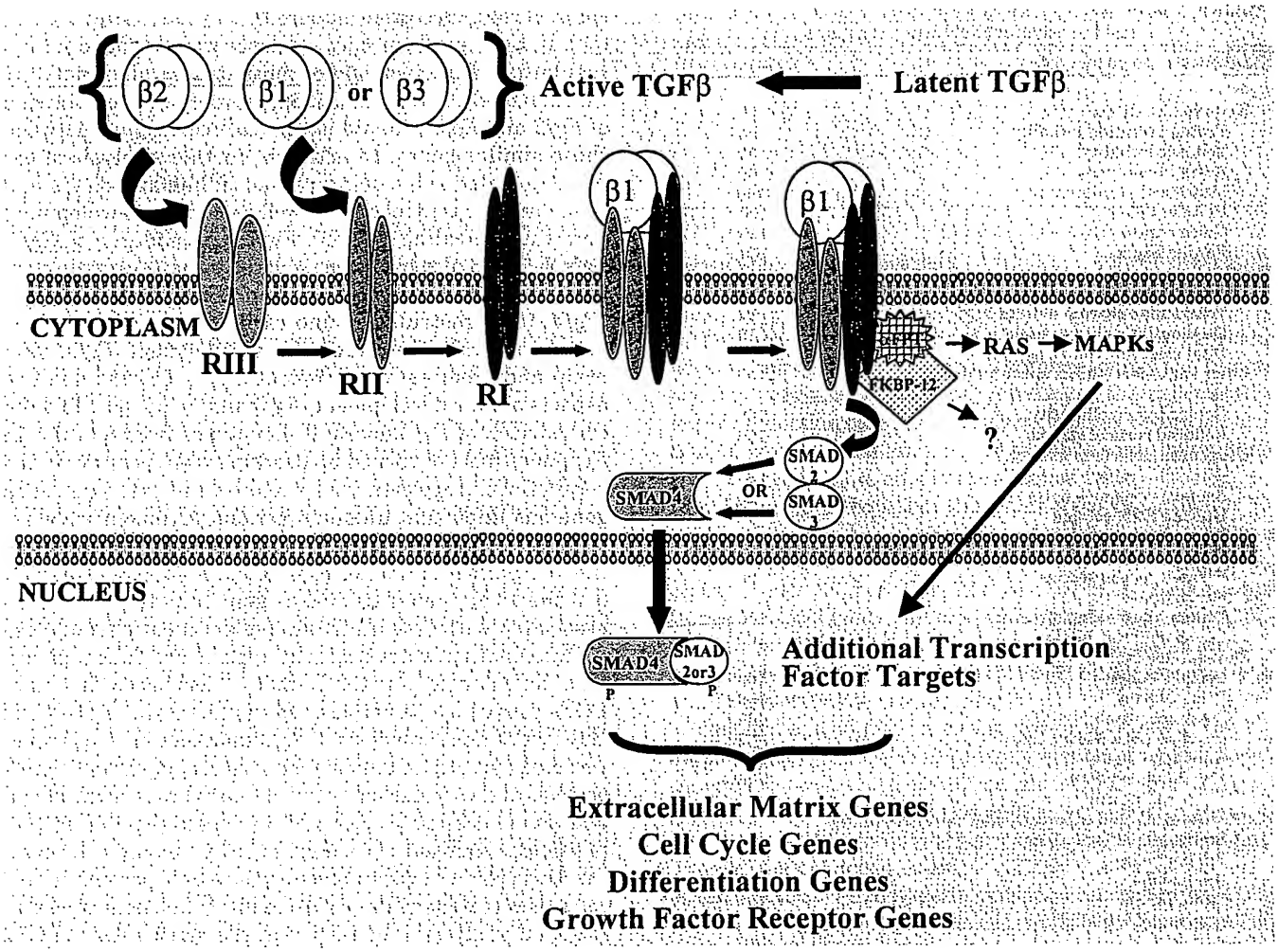


Figure 1. TGFβ signaling pathway. The TGFβ ligands, TGFβ1 (β1), TGFβ2 (β2), and TGFβ3 (β3), exist primarily in a latent form in vivo and are activated by mechanisms not yet clear. In general, TGFβ2 interacts with a TGFβ type III receptor (RIII) before interaction with TGFβ type II (RII) and TGFβ type I (RI) receptors; whereas, the TGFβ1 and TGFβ3 ligands can interact directly with the type II receptor. The ligand receptor complexes can then associate with several cytoplasmic molecules, farnesyl protein transferase (FPT) and FK506 binding protein-12 (FKBP-12), being two potential examples. The receptor-ligand complex signals to the nucleus through threonine/serine phosphorylation of a series of SMAD proteins (related to the Drosophila "mothers against decapentaplegic" protein) which then elicit transciptional regulation of extacellular matrix, cell cycle, differentiation and growth factor receptor genes. The roles of the associated cytoplasmic molecules FPT and FKBP-12 are not clear but are thought to involve RAS pathway signaling and modulation of signaling through the SMAD proteins.

the case for two of the myogenic genes known to be essential for specification of vertebrate skeletal muscle, *Myod* and *Myf5*. Even though the individual knockouts have muscle, and only the combined knockouts do not have muscle (45), it is now clear that each gene functions in the specification of distinct muscle cell lineages. Consequently, in the absence of one source of muscle cells, the other source may compensate for that (46, 47). This should be termed developmental compensation, rather than gene redundancy. On the other hand, with respect to retinoic acid receptors, there is also good evidence for functional redundancy. Similar to the myogenic genes, retinoic acid receptor gene knockout mice have few phenotypes, whereas the combined knockouts have many phenotypes (48, 49). Whether this turns out to be gene redundancy or another case of developmental compensation remains to be determined.

Lack of phenotype: As is the case for TGFβ, there also is a multitude of reports indicating that the FGFs 1 and 2 have important roles in numerous cell types and tissues. Consequently, when the *Fgf2* gene was knocked out by gene targeting, it was quite surprising that there was no obvious phenotype (50). The *Fgf2*$^{-/-}$ animals live a long, healthy life, and fertility and fecundity are normal. Even the pituitary gland, which is the best source of FGF2, appears not to have morphologic defects. The only evidence for any developmental abnormalities is found in hematopoiesis (50), where blood platelet counts are high, and in the cerebral cortex (51, 52), where morphometric analysis reveals decreased cell density. Clearly, these abnormalities are minor, compared with expectations. This was all the more evident because our transgenic mice, in which the human *FGF2* gene was ubiquitously overexpressed by the phosphoglycerate kinase pro-

Table 2. Nonoverlapping phenotypes of *Tgfb1*, *2*, and *3* knockout mice and the penetrance of those phenotypes

Knockout mouse phenotype	Penetrance (%)[a]
Tgfb1	
Embryo lethalities	50
Preimplantation lethality	50
Yolk sac lethality	50[b]
Adult phenotypes	50
Multifocal autoimmunity	100[c]
Platelet defect	100[c,d]
Colon cancer	100[c,d]
Failing heart	100[c,d]
Tgfb2 (all perinatal lethalities)	
Heart defects	100
Ventricular septum defects	94
Dual outlet right ventricle	19
Dual inlet left ventricle	25
Inner ear defect—lacks spiral limbus	100
Eyes	
Ocular hypercellularity	100
Reduced corneal stroma	100
Urogenital defects in kidney	
Dilated renal pelvis	30
Agenesis (females only)	20
Uterine horn ectopia	40
Testicular ectopia	100
Testis hypoplasia	20
Vas deferens dysgenesis	20
Lung-postnatal	
Dilated conducting airways	100
Collapsed bronchioles	100
Skeletal defects	
Occipital bone	100
Parietal bone	100
Squamous bone	100
Palatine bone (cleft palate)	22
Alisphenoid bone	100
Mandibular defects	100
Short radius and ulna	100
Missing deltoid tuberosity and third trochanter	94
Sternum malformations	25
Rib barreling	94
Rib fusions	13
Spina bifida	100
Tgfb3 (perinatal lethality)	
Cleft palate	100

[a] *See* Table 3 for background dependency of *Tgfb1* knockout phenotypes.
[b] Described in references 64, 67.
[c] Refers to percentage penetrance among animals that survive to birth.
[d] Unpublished observations.
Details on the remaining phenotypes can be found in the text and in references 29–31, 63.

moter (53), had very short legs, suggesting an important role of FGF2 in bone development, yet the bones of the knockout animals were normal. This apparent discrepancy between the transgenic and knockout mice indicates that some other FGF signals through the same FGF receptor as does FGF2, and that this other FGF is the true ligand that is important in bone development. Another possibility is that there is "developmental compensation" by alternative mechanisms. In other words, the absence of FGF2 may cause developmental abnormalities during bone development that are then compensated for by another developmental pathway. This alternative would not necessarily require a different FGF to be involved.

After we had made our first analysis of the *Fgf2* knockout mouse and did not find an obvious phenotype, it was easy to explain the "lack of phenotype" by invoking redundancy because there are at least 18 known *Fgf* genes. But in hindsight, it now appears more likely that all members of this large gene family have specific functions, even though they signal through receptors encoded by only four receptor genes (54). In *Fgf2* knockout mice, evidence was not found for upregulation of the two ligands most structurally related to FGF2, namely, FGFs 1 and 5 (50). Also, genetic combination of *Fgf2* and *Fgf5* (50) did not reveal redundancy between these similar genes. In addition, further analysis of the mice revealed roles being played in hematopoiesis and vascular tone control (50) as well as in brain development and wound healing (51, 52). Finally, in addition to *Fgf2*, *Fgfs* 3–5, 7, 8 also have been ablated by gene targeting, revealing functions in proliferation of the inner cell mass (*Fgf4*) (55); gastrulation and cardiac, craniofacial, forebrain, midbrain, and cerebellar development (*Fgf8*) (56); brain and inner ear development (*Fgf3*) (57, 58); and two aspects of hair development (*Fgf5* and 7) (59, 60). To date, comparison of *Fgf* knockout phenotypes from 6 of the 18 *Fgf* genes has not turned up overlap except possibly in the cerebellum. Together, these results indicate that each gene has important unique functions. Although a few redundant functions may eventually be found on combination of *Fgf2* with all other *Fgfs* except *Fgf5*, it is clear that 6 of the 18 *Fgf* genes studied by gene targeting have been associated with essentially unique knockout phenotypes.

To summarize, what originally appeared as "lack of phenotype" led many of us to the premature conclusion that other FGFs must have functions redundant to those of FGF2. However, further analysis of *Fgf2* knockout mice has since revealed a wealth of unique functions ranging from thrombocytosis and vascular tone control (50) to brain development and wound healing (51, 52). It is my expectation that further physiologic analysis of the *Fgf2* knockout mouse will reveal functions in the hypertrophic response to hypertension and responses to ischemia/reperfusion injury and bone injury. In the final analysis, it is likely that the major roles of FGF2 may have less to do with getting us to birth than with keeping us alive after birth, whereas several other FGFs clearly have developmental roles.

Effects of genetic background on phenotypic variation: From 100 years of mouse genetics, it has become clear that genetic background plays an important role in the susceptibility of mice to many disorders. Therefore, the phenotypes of knockout mouse strains will also have genetic background dependencies, as was first documented by the Magnuson and Wagner groups (61, 62). The *Tgfb1* knockout mice are an exceptional case in point (Table 3). On a mixed (50:50) 129 x CF1 background (CF1 is a partially outbred strain), about half of *Tgfb1* knockout mice die from a preimplantation developmental defect (63), and the other half die of an autoimmune-like multifocal inflammatory disease at about weaning age (29). If the targeted *Tgfb1* allele is backcrossed onto a C57BL/6 background, 99% of all knockout animals die of the preimplantation defect(63). However, if a *Tgfb1* knockout allele is put onto a mixed 129 x NIH/Ola x C57BL/6 background, embryo lethality is observed during yolk sac development, not during preimplantation development (64). With respect to the multifocal inflammatory disorder of *Tgfb1* knockout mice, if the targeted allele is put onto a 129 x CF1 mixed background (50:50), severe inflammation exists only in the stomach (29); on the mixed 129 x

Table 3. Background dependency of *Tgfb1* knockout phenotypes

Phenotype	Phenotype penetrance on various strains (%)						
	129 x CF1	129 x C57	129 x C3H	C57	129	C3H	129xC57x NIH/Ola
Preimplantation lethality	50	ND	ND	99	ND	ND	0
Yolk sac lethality[a]	0	0	ND	ND	ND	ND	50
Autoimmune disease	50	50	50	1	ND	ND	50
Gastric inflammation	90[b]	20[b]	ND	ND	ND	ND	ND
Intestinal inflammation	0	70[b]	ND	ND	ND	ND	ND
Colon cancer[c]	ND	ND	ND	ND	100	0	ND

Percentage of knockout animals of a given strain that have the designated phenotype.
[a]For details, *see* references 64, 67.
[b]Approximately 10% of animals with autoimmune disease have no detectable gastrointestinal tract inflammation.
[c]Unpublished observations.
ND = not determined.

NIH/Ola x C57BL/6 background, the intestines are more severely inflamed than is the stomach (65). Finally, on a predominantly 129 background (129 x CF1; ~97:3), *Tgfb1* knockout mice develop colon cancer if the inflammatory disorder can be eliminated by other genetic manipulations that render the mice immunodeficient (unpublished observations). However, on a predominantly C3H background, immunodeficient *Tgfb1* knockout mice do not develop colon cancer (66). These results suggest that modifier genes exist that can significantly affect the function of TGFβ1 in preimplantation development, yolk sac development, bowel and gastric inflammation, and colon tumor suppression. Progress toward localizing a modifier gene for the yolk sac developmental problem has been made (67).

What is the best genetic background for knockout mice? Because background-dependent phenotypic variability will likely be found for most knockout mice, it will be useful to backcross a targeted allele onto several mouse backgrounds to make congenic strains. In this section, it will be argued that putting a targeted allele on a mixed strain background will also provide useful information. This is not to say that congenic strains are not useful. Rather, the point to be made here is that there also are benefits to looking at mixed strain backgrounds. Again, our experience with *Tgfb* knockout mice will be instructive.

Generating homozygous mutant knockout animals on a mixed genetic background is faster. The ES cells are nearly always from a 129 strain, and the blastocysts into which the targeted ES cells are injected are nearly always C57BL/6. For reasons unknown, this is a good combination for establishing germline transmission of the injected ES cells. The resulting chimeras can then be crossed with any strain desired, but 129, C57BL/6, or Black Swiss mice are most often used, and CF1 mice were used in the case of our *Tgfb1* knockout mice. Heterozygous offspring from this crossing will then be inbred 129 or F1 hybrids of 129 and one of the other strains. Clearly then, the quickest route to having the knockout allele on an inbred strain is through 129. For the other strains several generations of backcrossing is required, which can take well over a year. Unfortunately, strain-129 mice have low fertility and fecundity. Consequently, the number of offspring per litter is usually fewer than six. Although 129 x C57BL/6 hybrids are more robust, upon backcrossing onto C57BL/6, litter size decreases. Tb the contrary, the Black Swiss and CF1 strains are robust, and litter size often is in excess of 12. The reason for this is probably because they are not truly inbred strains, but rather are partially outbred through random breeding within their respective strains. Therefore, one of the choices one has is to stay with "pure" genetics at the expense of a lower production rate and considerable delay before generation of experimental animals, or sacrifice some genetic purity to obtain a more efficient production colony. Ideally, one would want to do both, but this often is too expensive.

Mixed genetic background knockout mice often have a wider range of phenotypes. The *Tgfb1* knockout mice backcrossed onto either the 129 or C57BL/6 background (congenics) yield only embryo lethality (63, unpublished observations). On the other hand, when the knockout allele is maintained on mixed genetic backgrounds, embryo and adult phenotypes are maintained.

The *Tgfb2* & *Tgfb3* knockout mice provide further examples. The *Tgfb2* knockout mice have more than two dozen congenital defects and die either immediately preceding or during birth, or within 2 h thereafter (30). Table 2 indicates that most of the phenotypes are only partially penetrant. Though it is not documented, it is likely that the penetrance of some of these phenotypes would increase to nearly 100%, and some of the other phenotypes would disappear were we to put the *Tgfb2* knockout allele on inbred backgrounds. Hence, the mixed strain background probably provided more information than would congenic strains.

The *Tgfb3* knockout mice have a cleft palate (31). One colony of *Tgfb3* knockout mice was left as a mixed background (129 x CF1; 50:50) strain, whereas another colony was backcrossed several generations to the C57BL/6 strain. These two colonies had considerable expressivity differences; the inbred colony had more severe clefting than did the mixed background colony. In the latter, expressivity of clefting varied widely from animal to animal. This variable expressivity within the mixed background colony provided us with the opportunity to obtain far more data on development of the cleft palate and was, therefore, more useful for making assumptions about the cellular and molecular mechanisms by which TGFβ3 supports palate fusion. Hence, using the *Tgfb3* knockout mice, the mixed strain background provided more information than did the congenic strain. Consequently, a wider range of penetrance and expressivity of phenotype is a major advantage of investigating knockout phenotypes in mixed background colonies. Further, variable penetrance of phenotype in a mixed background colony suggests that there are modifier genes for each phenotype that could be obtained by linkage studies.

Conclusions

Questions have been addressed that arose from the last 8 years in which knockout mice have been investigated to analyze gene function at the whole animal level. These questions concern gene redundancy, apparent lack of phenotype in a surprising number of knockout strains, and effects of genetic background on knockout phenotype. Using data obtained principally from *Tgfb* and *Fgf* knockout mice, it is argued that there is probably little redundancy in the genome (i.e., that few genes are dispensable for survival of the species). Apparent lack of phenotype more likely reflects our inability to ask the right questions, or our lack of tools to answer them, than it does a true lack of function. Finally, discussion of genetic background phenotype variability, including variable penetrance and expressivity, was used to present some of the advantages of working with mixed genetic background colonies of knockout mice. For all the examples given here, there are counter examples that must be taken seriously; consequently, these arguments must not be taken as absolutes. For example, if a gene in a particular mouse strain has recently been duplicated, it will most likely be redundant. If one is studying tissue rejection in a knockout mouse, the genetic background obviously must be well defined and preferably inbred. Or, if one wants to use the susceptibility of a particular mouse strain to cancer to investigate the function of the knockout gene in progression of that cancer, the knockout allele must be put on that mouse strain. In general, however, when setting up approaches for investigating a new gene knockout mouse, I believe one would be well advised to assume that: there is little gene redundancy in mammals; there are knockout phenotypes even if none are immediately apparent; and investigating phenotypes in mixed genetic background colonies may not only reveal more phenotypes, but may lead to better understanding of the molecular or cellular mechanism underlying the phenotype, and may lead to modifier gene discovery.

Acknowledgements

I thank all the members of the Doetschman lab for their intellectual stimulation, for the constant generation of exciting experimental results, and for their ability to become intrigued rather than discouraged when those results are not what was expected.

The work discussed here was supported by grant nos. HD26471, HL58511, HL41496, ES05652, ES06096, and AR44059 to T. D.

References

1. **Mansour, S. L.** 1990. Gene targeting in murine embryonic stem cells: introduction of specific alterations into the mammalian genome. Genet. Anal. Tech. Appl. **7**:219–227.
2. **Koller, B. H., and O. Smithies.** 1992. Altering genes in animals by gene targeting. Ann. Rev. Immunol. **10**:705–730.
3. **Bradley, A., R. Ramirez-Solis, H. Zheng, *et al.*** 1992. Genetic manipulation of the mouse via gene targeting in embryonic stem cells. Ciba Found. Symp. **165**:256–269.
4. **Doetschman, T.** 1994. Gene transfer in embryonic stem cells, p. 115–146. *In* C. A. Pinkert (ed.), Transgenic animal technology: a laboratory handbook. Academic Press, Inc., New York.
5. **Jaenisch, R.** 1988. Transgenic animals. Science **240**: 1468–1474.
6. **Hanahan, D.** 1989. Transgenic mice as probes into complex systems. Science **246**:1265–1275.
7. **Doetschman, T., P. Williams, and N. Maeda.** 1988. Establishment of hamster blastocyst-derived embryonic stem (ES) cells. Dev. Biol. **127**:224–227.
8. **Iannaccone, P. M., G. U. Taborn, R. L. Garton, *et al.*** 1994. Pluripotent embryonic stem cells from the rat are capable of producing chimeras. Dev. Biol. **163**:288–292.
9. **Graves, K. H., and R. W. Moreadith.** 1993. Derivation and characterization of putative pluripotential embryonic stem cells from preimplantation rabbit embryos. Mol. Reprod. Dev. **36**:424–433.
10. **Schoonjans, L., G. M. Albright, J. L. Li, *et al.*** 1996. Pluripotential rabbit embryonic stem (ES) cells are capable of forming overt coat color chimeras following injection into blastocysts. Mol. Reprod. Dev. **45**:439–443.
11. **Wheeler, M. B.** 1994. Development and validation of swine embryonic stem cells: a review. Reprod. Fertil. Dev. **6**:563–568.
12. **Shim, H., A. Gutierrez-Adan, L. R. Chen, *et al.*** 1997. Isolation of pluripotent stem cells from cultured porcine primordial germ cells. Biol. Reprod. **57**:1089–1095.
13. **Piedrahita, J. A., K. Moore, B. Oetama, *et al.*** 1998. Generation of transgenic porcine chimeras using primordial germ cell-derived colonies. Biol. Reprod. **58**:1321–1329.
14. **Cherny, R. A., T. M. Stokes, J. Merei, et al.** 1994. Strategies for the isolation and characterization of bovine embryonic stem cells. Reprod. Fertil. Dev. **6**:569–575.
15. **First, N. L., M. M. Sims, S. P. Park, *et al.*** 1994. Systems for production of calves from cultured bovine embryonic cells. Reprod. Fertil. Dev. **6**:553–562.
16. **Sun, L., C. S. Bradford, and D. W. Barnes.** 1995. Feeder cell cultures for zebrafish embryonal cells in vitro. Mol. Mar. Biol. Biotechnol. **4**:43–50.
17. **Valancius, V., and O. Smithies.** 1991. Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. Mol. Cell. Biol. **11**:1402–1408.
18. **Hasty, P., R. Ramirez-Solis, R. Krumlauf, *et al.*** 1991. Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. Nature **350**:243–246.
19. **Askew, G. R., T. Doetschman, and J. B. Lingrel.** 1993. Site- directed point mutations in embryonic stem cells: a gene-targeting tag-and-exchange strategy. Mol. Cell. Biol. **13**: 4115–4124.
20. **Gu, H., Y. R. Zou, and K. Rajewsky.** 1993. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell **73**:1155–1164.
21. **Gu, H., J. D. Marth, P. C. Orban, *et al.*** 1994. Deletion of a DNA polymerase *beta* gene segment in T cells using cell type-specific gene targeting. Science **265**:103–106.
22. **Schwenk, F., U. Baron, and K. Rajewsky.** 1995. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. Nucleic Acids Res. **23**:5080–5081.
23. **Fishman, G. I., M. L. Kaplan, and P. M. Buttrick.** 1994. Tetracycline-regulated cardiac gene expression in vivo. J. Clin. Invest. **93**:1864–1868.
24. **Shockett, P., M. Difilippantonio, N. Hellman, *et al.*** 1995. A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. Proc. Natl. Acad. Sci. USA **92**:6522–6526.
25. **Kuhn, R., F. Schwenk, M. Aguet, *et al.*** 1995. Inducible gene targeting in mice. Science **269**:1427–1429.
26. **No, D., T. P. Yao, and R. M. Evans.** 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. Proc. Natl. Acad. Sci. USA **93**:3346–3351.
27. **Pelton, R. W., B. L. Hogan, D. A. Miller, *et al.*** 1990. Differential expression of genes encoding TGFs *beta* 1, *beta* 2, and *beta* 3 during murine palate formation. Dev. Biol. **141**: 456–460.
28. **Pelton, R. W., B. Saxena, M. Jones, *et al.*** 1991. Immunohistochemical localization of TGF *beta* 1, TGF *beta* 2, and TGF *beta* 3 in the mouse embryo: expression patterns suggest multiple roles during embryonic development. J. Cell Biol. **115**:1091–1105.

29. **Shull, M. M., I. Ormsby, A. B. Kier, *et al.*** 1992. Targeted disruption of the mouse transforming growth factor-*beta* 1 gene results in multifocal inflammatory disease. Nature **359:** 693–699.
30. **Sanford, L. P., I. Ormsby, G. A. Gittenberger-de, *et al.*** 1997. TGF *beta* 2 knockout mice have multiple developmental defects that are non-overlapping with other TGF *beta* knockout phenotypes. Development **124:**2659–2670.
31. **Proetzel, G., S. A. Pawlowski, M. V. Wiles, *et al.*** 1995. Transforming growth factor-*beta* 3 is required for secondary palate fusion. Nat. Genet. **11:**409–414.
32. **Flaumenhaft, R., M. Abe, Y. Sato, *et al.*** 1993. Role of the latent TGF-*beta* binding protein in the activation of latent TGF-*beta* by co-cultures of endothelial and smooth muscle cells. J. Cell Biol. **120:**995–1002.
33. **Munger, J. S., J. G. Harpel, F. G. Giancotti, *et al.*** 1998. Interactions between growth factors and integrins: latent forms of transforming growth factor-*beta* are ligands for the integrin *alpha*$_v$ *beta*$_1$. Mol. Biol. Cell **9:**2627–2638.
34. **Munger, J. S., J. G. Harpel, P. E. Gleizes, *et al.*** 1997. Latent transforming growth factor-*beta*: structural features and mechanisms of activation. Kidney Int. **51:**1376–1382.
35. **Nunes, I., P. E. Gleizes, C. N. Metz, *et al.*** 1997. Latent transforming growth factor-*beta* binding protein domains involved in activation and transglutaminase-dependent crosslinking of latent transforming growth factor-*beta*. J. Cell Biol. **136:**1151–1163.
36. **Cheifetz, S., T. Bellon, C. Cales, *et al.*** 1992. Endoglin is a component of the transforming growth factor-*beta* receptor system in human endothelial cells. J. Biol. Chem. **267:** 19027–19030.
37. **Lamarre, J., J. Vasudevan, and S. L. Gonias.** 1994. Plasmin cleaves betaglycan and releases a 60 kDa transforming growth factor-*beta* complex from the cell surface. Biochem. J. **302:**199–205.
38. **Sankar, S., N. Mahooti-Brooks, M. Centrella, *et al.*** 1995. Expression of transforming growth factor type III receptor in vascular endothelial cells increases their responsiveness to transforming growth factor *beta* 2. J. Biol. Chem. **270:**13567–13572.
39. **Heldin, C. H., K. Miyazono, and P. ten Dijke.** 1997. TGF-*beta* signalling from cell membrane to nucleus through SMAD proteins. Nature **390:**465–471.
40. **Kretzschmar, M., and J. Massague.** 1998. SMADs: mediators and regulators of TGF-*beta* signaling. Curr. Opin. Genet. Dev. **8:**103–111.
41. **Bruno, E., S. K. Horrigan, D. Van Den Berg, *et al.*** 1998. The Smad5 gene is involved in the intracellular signaling pathways that mediate the inhibitory effects of transforming growth factor-*beta* on human hematopoiesis. Blood **91:** 1917–1923.
42. **Afrakhte, M., A. Moren, S. Jossan, *et al.*** 1998. Induction of inhibitory Smad6 and Smad7 mRNA by TGF-*beta* family members. Biochem. Biophys. Res. Commun. **249:**505–511.
43. **Nakayama, T., H. Gardner, L. K. Berg, *et al.*** 1998. Smad6 functions as an intracellular antagonist of some TGF-*beta* family members during Xenopus embryogenesis. Genes Cells **3:**387–394.
44. **Itoh, S., M. Landstrom, A. Hermansson, *et al.*** 1998. Transforming growth factor *beta* 1 induces nuclear export of inhibitory Smad7. J. Biol. Chem. **273:**29195–29201.
45. **Rudnicki, M. A., P. N. Schnegelsberg, R. H. Stead, *et al.*** 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. Cell **75:**1351–1359.
46. **Kablar, B., K. Krastel, C. Ying, *et al.*** 1997. MyoD and Myf-5 differentially regulate the development of limb versus trunk skeletal muscle. Development **124:**4729–4738.
47. **Ordahl, C. P., and B. A. Williams.** 1998. Knowing chops from chuck: roasting myoD redundancy. Bioessays **20:**357–362.
48. **Lohnes, D., M. Mark, C. Mendelsohn, *et al.*** 1994. Function of the retinoic acid receptors (RARs) during development (I). Craniofacial and skeletal abnormalities in RAR double mutants. Development **120:**2723–2748.
49. **Mendelsohn, C., D. Lohnes, D. Decimo, *et al.*** 1994. Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants. Development **120:**2749–2771.
50. **Zhou, M., R. L. Sutliff, R. J. Paul, *et al.*** 1998. Fibroblast growth factor 2 control of vascular tone. Nat. Med. **4:**201–207.
51. **Ortega, S., M. Ittmann, S. H. Tsang, *et al.*** 1998. Neuronal defects and delayed wound healing in mice lacking fibroblast growth factor 2. Proc. Natl. Acad. Sci. USA **95:**5672–5677.
52. **Dono, R., G. Texido, R. Dussel, *et al.*** 1998. Impaired cerebral cortex development and blood pressure regulation in FGF-2-deficient mice. EMBO J. **17:**4213–4225.
53. **Coffin, J. D., R. Z. Florkiewicz, J. Neumann, *et al.*** 1995. Abnormal bone growth and selective translational regulation in basic fibroblast growth factor (FGF-2) transgenic mice. Mol. Biol. Cell **6:**1861–1873.
54. **Ornitz, D. M., J. Xu, J. S. Colvin, *et al.*** 1996. Receptor specificity of the fibroblast growth factor family. J. Biol. Chem. **271:**15292–15297.
55. **Feldman, B., W. Poueymirou, V. E. Papaioannou, *et al.*** 1995. Requirement of FGF-4 for postimplantation mouse development. Science **267:**246–249.
56. **Meyers, E. N., M. Lewandoski, and G. R. Martin.** 1998. An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. Nat. Genet. **18:**136–141.
57. **Mansour, S. L., J. M. Goddard, and M. R. Capecchi.** 1993. Mice homozygous for a targeted disruption of the proto-oncogene int-2 have developmental defects in the tail and inner ear. Development **117:**13–28.
58. **McKay, I. J., J. Lewis, and A. Lumsden.** 1996. The role of FGF-3 in early inner ear development: an analysis in normal and kreisler mutant mice. Dev. Biol **174:**370–378.
59. **Hebert, J. M., T. Rosenquist, J. Gotz, *et al.*** 1994. FGF5 as a regulator of the hair growth cycle: evidence from targeted and spontaneous mutations. Cell **78:**1017–1025.
60. **Guo, L., L. Degenstein, and E. Fuchs.** 1996. Keratinocyte growth factor is required for hair development but not for wound healing. Genes Dev. **10:**165–175.
61. **Threadgill, D. W., A. A. Dlugosz, L. A. Hansen, *et al.*** 1995. Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science **269:**230–234.
62. **Sibilia, M., and E. F. Wagner.** 1995. Strain-dependent epithelial defects in mice lacking the EGF receptor. Science **269:**234–238.
63. **Kallapur, S., I. Ormsby, and T. Doetschman.** 1999. Strain dependency of TGFβ1 function during embryogenesis. Mol. Reprod. Develop. **52:**341–349.
64. **Dickson, M. C., J. S. Martin, F. M. Cousins, *et al.*** 1995. Defective haematopoiesis and vasculogenesis in transforming growth factor-*beta* 1 knock out mice. Development **121:**1845–1854.
65. **Kulkarni, A. B., C. G. Huh, D. Becker, *et al.*** 1993. Transforming growth factor *beta* 1 null mutation in mice causes excessive inflammatory response and early death. Proc. Natl. Acad. Sci. USA **90:**770–774.
66. **Diebold, R. J., M. J. Eis, M. Yin, *et al.*** 1995. Early-onset multifocal inflammation in the transforming growth factor *beta* 1-null mouse is lymphocyte mediated. Proc. Natl. Acad. Sci. USA **92:**12215–12219.
67. **Bonyadi, M., S. A. Rusholme, F. M. Cousins, *et al.*** 1997. Mapping of a major genetic modifier of embryonic lethality in TGF *beta* 1 knockout mice. Nat. Genet. **15:**207–211.

EXHIBIT I

Researchers to Gain Wider Access to Knockout Mice
<<http://www.genome.gov/17015131>>

See Also:

Knockout Mouse Project RFAs

Knockout Mice
Fact Sheet

U.S. Department of Health and Human Services
NIH News
National Institutes of Health

Researchers to Gain Wider Access to Knockout Mice

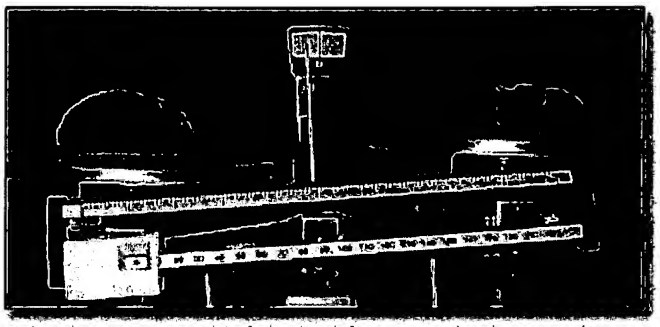
Trans-NIH Effort Provides New Models for Understanding Human Disease

BETHESDA, Md., Wed., Oct. 5, 2005 - The National Institutes of Health (NIH) today announced contracts that will give researchers unprecedented access to two private collections of knockout mice, providing valuable models for the study of human disease and laying the groundwork for a public, genome-wide library of knockout mice.

Under terms of three-year contracts jointly funded by 19 NIH institutes, centers and offices, Deltagen Inc. of San Carlos, Calif., and Lexicon Genetics Incorporated of The Woodlands, Texas will provide NIH and its scientific partners with access to extensively characterized lines of mice in which a specific gene has been disrupted, or "knocked out." In the first year of the contract, NIH will expend about $10 million to acquire about 250 lines of knockout mice.

For each mouse line, the contractors will provide not only the mouse line itself, but also detailed, objective data on the impact of the specific gene deletion on the mouse's phenotype, which includes appearance, health, fitness, behavior, ability to reproduce, and radiological and microscopic data. Such comprehensive information on such a large group of mice has never been available to public sector researchers, and is expected to greatly accelerate efforts to explore gene functions in health and disease.

"Our decision to procure these knockout mouse lines and data and make them available to the research community will yield tremendous benefits, both in the short and long terms," said NIH Director Elias A. Zerhouni, M.D. "This trans-NIH initiative will place important mouse models into the hands of researchers, speeding advances in the understanding of human disease and the development of new therapies. It also represents a significant step in the direction of launching an international project to systematically knock out all genes in the mouse."



A knockout mouse model of obesity, left, compared with a normal mouse. Photo courtesy of Lexicon Genetics Incorporated.

Since the early 1980s, when recombinant DNA technology was used to create the first such animals, knockout mice have proven to be one of the most powerful tools available to study the function of genes and to create mouse models of human disease. Researchers have produced knockout mice with characteristics similar to humans suffering from a wide range of disorders, including cancer, heart disease, neurological disorders and even obesity.

The process used by NIH to select the mouse lines involved a rigorous scientific review process that evaluated information on the knocked out gene, the reliability of the method used to produce the knockout, and whether the mouse line possesses a "reporter" gene, which enables researchers to analyze the pattern of the knockout gene's expression in various mouse tissues.

"This is exciting news for all researchers working to understand the complex underpinnings of human biology in health and disease. Knockout mice provide one of the quickest, most cost effective ways to explore gene function. It is essential that we make it possible for more researchers to tap into this power," said James Battey, M.D., Ph.D., director of the National Institute on Deafness and Other Communication Disorders. Dr. Battey serves as chair of the Trans-NIH Mouse Initiative, which develops priorities for mouse genomics and genetic resources at NIH.

The new contracts provide NIH with irrevocable, perpetual, worldwide, royalty-free licenses to use and distribute to academic and non-profit researchers these lines of knockout mice. The mouse lines, which will be stored in the form of frozen embryos, frozen sperm and frozen embryonic stem (ES) cells, will be delivered to NIH-funded mouse repositories that supply mice to universities, medical schools and research labs all over the world. When researchers express interest in obtaining a certain knockout mouse line, the repositories will send them live mice, frozen embryos, sperm, and/or ES cells, so they can study the

mice in their own labs. All data on the mice will be made available to researchers worldwide without restriction in publicly available databases on the Web.

Under the license agreements with Deltagen and Lexicon, researchers who receive the knockout mice lines are free to publish any results from research involving the line and also to seek patent or other intellectual property protection for any of the inventions or discoveries resulting from such research.

The 19 NIH institutes, centers and offices contributing to the contracts are: National Center for Complementary and Alternative Medicine, National Center for Research Resources, National Eye Institute, National Human Genome Research Institute, National Heart, Lung and Blood Institute, National Institute on Aging, National Institute of Alcohol Abuse and Alcoholism, National Institute of Allergy and Infectious Diseases, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institute of Child Health and Human Development, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute on Drug Abuse, National Institute of Environmental Health Sciences, National Institute of General Medical Sciences, National Institute of Mental Health, National Institute of Neurological Disorders and Stroke, National Institute of Nursing Research, and the Office of AIDS Research.

In order to build upon the acquisition of knockout mice available from the private sector, the NIH in September issued a set of requests for applications to establish a Knockout Mouse Project. The ambitious goal of this trans-NIH program is to produce a comprehensive resource of mouse mutants in which every gene in the mouse genome has been knocked out. The resource will serve to further the value of the mouse as a powerful and important tool in the study of human health and disease.

For more information on what knockout mice are, how they are made and what they are used for, go to www.genome.gov/12514551. To download a high-resolution photo of a knockout mouse, go to www.genome.gov/17015128.

NIH is the steward of medical and behavioral research for the United States. It is an agency of the U.S. Department of Health and Human Services. For more information about NIH, go to www.nih.gov.

Contact:

Geoff Spencer, NHGRI
(301) 402-0911
spencerg@mail.nih.gov

Top of page

Posted: October 2005

EXHIBIT J

Deltagen and Merck Enter Into DeltaBase License Agreement
<<http://www.sec.gov/Archives/edgar/data/1034072/000089843002000648/dex992.htm>>

EX-99.2 4 dex992.htm DELTAGEN, INC. NEWS RELEASE DTD FEBRUARY 11, 2002

Exhibit 99.2

Deltagen and Merck Enter Into DeltaBase License Agreement

REDWOOD CITY, Calif., Feb. 11/PRNewswire-FirstCall/ — Deltagen, Inc. (Nasdaq: DGEN) announced today that it has entered into a license agreement to provide Merck & Co., Inc. (NYSE: MRK) with access to Deltagen's proprietary DeltaBase™ product, a powerful resource tool for the understanding of in vivo mammalian gene function information.

Merck will have non-exclusive access to information related to 750 genes selected for their biological interest that have been functionally characterized and entered into DeltaBase. Merck will also have access to certain of the corresponding DeltaBase intellectual property rights. Financial terms were not disclosed. "We are delighted to extend our relationship with Merck, which has a worldwide reputation in pharmaceutical discovery and development, by adding them as a DeltaBase subscriber," said William Matthews, Ph.D., president and chief executive officer at Deltagen.

Through its proprietary product DeltaBase™, Deltagen provides pharmaceutical companies with critical information to better understand the in vivo function of mammalian genes, their relationship to other genes and the biochemical pathways for large segments of the genome. Each gene is the focus of an exhaustive investigation; more than 20,000 pieces of data from each gene are distilled into disease-relevant frameworks that include proprietary knockout mouse phenotypic, expression profile data and other proprietary target validation data. Information in DeltaBase is generated using Deltagen's large-scale mouse gene knockout technology and standardized phenotypic analysis protocols.

Deltagen is a genomic-based biotechnology company headquartered in Redwood City, California, that provides data to pharmaceutical and biotechnology companies on the function, role and disease relevance of mammalian genes. This information may facilitate the discovery and validation of drug targets to advance the development of new genomic-based medicines. Deltagen's principal product, DeltaBase, provides a database of in vivo derived, mammalian gene function information. In addition, the company is dedicated to determine the function of secreted proteins and is undertaking the discovery and development of biotechnology drug candidates internally or in collaboration with other parties. Current DeltaBase subscribers can be found on Deltagen's website, *www.deltagen.com.*

Except for the historical information contained herein, the matters set forth in this press release, including statements as to the role that Deltagen's DeltaBase product and gene function database information will play in third-party research programs and the extent to which genome-based research will assist researchers in their drug discovery efforts, are forward-looking statements within the meaning of the "safe harbor" provisions of the

Private Securities Litigation Reform Act of 1995. These forward-looking statements are subject to risks and uncertainties that may cause actual results to differ materially from those set forth in the forward-looking statements, including the extent to which genomic databases are utilized in pharmaceutical research and development; the ability of Deltagen to provide products and services that meet market needs; the impact of competition and alternative technologies, processes and approaches; and other risks cited in the risk factors sections of Deltagen's Annual Report on Form 10-K filed with the Securities and Exchange Commission and Deltagen's other securities filings with the Commission. These forward-looking statements speak only as of the date hereof. Deltagen disclaims any intent or obligation to update these forward-looking statements.

EXHIBIT K

Declaration of Robert Driscoll, filed June 22, 2005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Keith D. Allen	Examiner:	Qian, Celine X
Serial No.:	10/005,467	Group Art Unit:	1632
Filed:	December 4, 2001	Docket No.:	R758/75658.029500
Title:	Transgenic Mice Containing PTP36 Tyrosine Phosphatase Gene Disruptions		

DECLARATION OF ROBERT DRISCOLL PURSUANT TO 37 C.F.R. § 1.132

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

I, Robert Driscoll, residing at 23 Chicory Lane, San Carlos, CA 94070, hereby declare:

1. I am presently employed as Vice President of Intellectual Property & Legal Affairs at Assignee, Deltagen, Inc., in San Carlos, CA. I have also previously served as the Company's Senior Director of Intellectual Property, in which position I managed and oversaw the Company's intellectual property portfolio, including the Company's patent filings. I possess a Ph.D in Chemistry, received from the California Institute of Technology. I also possess a J.D., received from Loyola Law School, Los Angeles. I am a registered patent attorney (Reg. No. 47,536).

2. I am familiar with the above-cited application. I am familiar with the Office Action mailed March 22, 2005. I am aware that the Examiner has rejected the claims, in part, for allegedly failing to meet the utility requirement. I am also aware that the Applicant has argued that a commercial sale of a mouse with a disrupted PTP36 tyrosine phosphatase allele within the scope of the claimed subject matter ("PTP36 tyrosine phosphatase gene knockout mouse") should satisfy the utility requirement.

3. In support of the Applicant's aforementioned argument, I hereby state that I have reviewed Deltagen's internal sales records regarding the PTP36 tyrosine phosphatase gene

knockout mouse. According to these records, the PTP36 tyrosine phosphatase gene knockout mouse has been delivered to at least one (1) large pharmaceutical company. The contractual terms by which the mice were transferred prohibit Deltagen from identifying the name of this company. However, the company is ranked among the top 10 pharmaceutical companies worldwide (based on sales).

4. It is my understanding, based on communications with our pharmaceutical company customers, that transgenic knockout mice obtained from Deltagen are used for studying gene function and for human therapeutic drug development.

5. I further declare that all statements made herein of my own knowledge are true; and further that these statements were made with knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the above-referenced application or any patent issuing thereon.

Robert Driscoll, Ph.D, Reg. No. 47,536

14 June 2005

Date

EXHIBIT L

DiMiasi et al., Journal of Health Economics 22 (2003) 151–185

Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Health Economics 22 (2003) 151–185
JOURNAL OF HEALTH ECONOMICS
www.elsevier.com/locate/econbase

The price of innovation: new estimates of drug development costs

Joseph A. DiMasi [a,*], Ronald W. Hansen [b], Henry G. Grabowski [c]

[a] *Tufts Center for the Study of Drug Development, Tufts University, 192 South Street, Suite 550, Boston, MA 02111, USA*
[b] *William E. Simon Graduate School of Business Administration, University of Rochester, Rochester, NY, USA*
[c] *Department of Economics, Duke University, Durham, NC, USA*

Received 17 January 2002; received in revised form 24 May 2002; accepted 28 October 2002

Abstract

The research and development costs of 68 randomly selected new drugs were obtained from a survey of 10 pharmaceutical firms. These data were used to estimate the average pre-tax cost of new drug development. The costs of compounds abandoned during testing were linked to the costs of compounds that obtained marketing approval. The estimated average out-of-pocket cost per new drug is US$ 403 million (2000 dollars). Capitalizing out-of-pocket costs to the point of marketing approval at a real discount rate of 11% yields a total pre-approval cost estimate of US$ 802 million (2000 dollars). When compared to the results of an earlier study with a similar methodology, total capitalized costs were shown to have increased at an annual rate of 7.4% above general price inflation.

JEL classification: L65; O31

Keywords: Innovation; R&D cost; Pharmaceutical industry; Discount rate; Technical success rates

1. Introduction

Innovations in the health sciences have resulted in dramatic changes in the ability to treat disease and improve the quality of life. Expenditures on pharmaceuticals have grown faster than other major components of the health care system since the late 1990s. Consequently, the debates on rising health care costs and the development of new medical technologies have focused increasingly on the pharmaceutical industry, which is both a major participant in the health care industry and a major source of advances in health care technologies.

* Corresponding author. Tel.: +1-617-636-2116.
E-mail address: joseph.dimasi@tufts.edu (J.A. DiMasi).

0167-6296/03/$ – see front matter
doi:10.1016/S0167-6296(02)00126-1

One of the key components of the discussion is the role of private sector pharmaceutical industry investments in R&D and an understanding of the factors that affect this process. Although the industry engages in many forms of innovation, in general the most significant is the discovery and development of new chemical and biopharmaceutical entities that become new therapies. Our prior research (DiMasi et al., 1991) found that the discovery and development of new drugs is a very lengthy and costly process. In the research-based drug industry, R&D decisions have very long-term ramifications, and the impact of market or public policy changes may not be fully realized for many years. From both a policy perspective, as well as an industrial perspective, it is therefore important to continue to analyze the components of and trends in the costs of pharmaceutical innovation.

In this paper we will build on research conducted by the current authors (DiMasi et al., 1991) and others on the economics of pharmaceutical R&D. As we described in our prior study, "Empirical analyses of the cost to discover and develop NCEs are interesting on several counts. First, knowledge of R&D costs is important for analyzing issues such as the returns on R&D investment. Second, the cost of a new drug has direct bearing on the organizational structure of innovation in pharmaceuticals. In this regard, higher real R&D costs have been cited as one of the main factors underlying the recent trend toward more mergers and industry consolidation. Third, R&D costs also influence the pattern of international resource allocation. Finally, the cost of R&D has become an important issue in its own right in the recent policy deliberations involving regulatory requirements and the economic performance of the pharmaceutical industry". In the decade that has followed the publication of our earlier study, these issues remain paramount. In addition, the congressional debates on Medicare prescription drug coverage and various new state initiatives to fill gaps in coverage for the elderly and the uninsured have intensified the interest in the performance of the pharmaceutical industry.

In the current study we are not attempting to directly answer the policy debates mentioned above. Rather, our focus is on providing new estimates of economic parameters associated with the drug development process. In particular, we concentrate on estimates of the costs of pharmaceutical innovation. Our prior estimates have been used by the Office of Technology assessment (OTA), the Congressional Budget Office (CBO), and various researchers to analyze policy questions such as the effects on R&D activities of health care financing reform or changes in intellectual property legislation related to the pharmaceutical industry.

The approach used in this paper follows our previous study (DiMasi et al., 1991) and the earlier work by Hansen (1979). Given the similarity in methodologies, we are able to compare our results in the current study with the estimates in the earlier studies to illustrate trends in development costs. All three studies used micro-level data on the cost and timing of development obtained through confidential surveys of pharmaceutical firms for a random sample of new drugs first investigated in humans by these firms. In the current study, the new drugs were first tested in humans anywhere in the world between 1983 and 1994. The reported development costs ran through 2000.Ultimately, we are interested in the expected cost of development per approved new drug. The uncertainties in the research and development process result in expenditures on many development projects that are not successful in producing a marketed product. However, to produce an estimate of expected cost for a marketed product, we must allocate the costs of the unsuccessful projects to those

that result in a marketed new product. The R&D process is lengthy, and as such it is important to know at what stage of development expenses occur. Viewed as an investment project, it is necessary to know both the amount of expenditures and the timing of these expenditures, since funds committed to R&D in advance of any returns from sales have both a direct and an opportunity cost. We used a unique database to estimate various cost parameters in the development process. Of particular concern is the estimation of the average pre-tax cost of new drug development, since we are interested in the resource costs of new drug development and how they have changed over time.

1.1. Previous studies of the cost of pharmaceutical innovation

A summary of early studies of the cost of drug development can be found in the authors' previous study (DiMasi et al., 1991) and in OTA (1993). In brief, the early studies were either based on a case study of a specific drug (usually ignoring the cost of failed projects) or relied on aggregate data. Since the R&D process often extends for a decade or more and the new drug development process often changes, it is difficult to estimate the cost of development from aggregated annual data. In contrast, the study by Hansen (1979) and the current authors' previous study (DiMasi et al., 1991) estimated development cost based on data supplied by firms for a representative sample of drug development efforts.

DiMasi et al. (1991) used data on self-originated new drugs to estimate the average cost of developing a new drug. They obtained data from 12 pharmaceutical firms on the research and development costs of 93 randomly selected new drugs that entered clinical trials between 1970 and 1982. From these data they estimated the average pre-tax out-of-pocket cost per approved drug to be US$ 114 million (1987 dollars). Since these expenditures were spread out over nearly a dozen years, they capitalized these expenditures to the date of marketing approval using a 9% discount rate. This yielded an estimate of US$ 231 million (1987 dollars). Measured in constant dollars, this value is more than double that obtained by Hansen for an earlier sample. DiMasi et al. (1991) also found that the average cost of the first two phases of clinical trials doubled between the first and second half of their sample. This led to the expectation that development costs would be higher in future samples.

Based on an analysis by Myers and Shyam-Sunder performed for the OTA, the OTA (1993) report noted that the cost-of-capital for the industry was roughly 10% in the early 1980s. This is moderately higher than the 9% used by DiMasi et al. (1991). The OTA also recalculated the DiMasi et al. (1991) numbers using an interest rate that varied over the life of the R&D cycle thereby raising the cost estimate by US$ 100 million in 1990 dollars.[1] The OTA presented both pre- and post-tax cost estimates.

[1] The OTA applied a range of discount rates that varied with the time to marketing approval. They chose 14% for the earliest stage R&D and 10% for development just prior to approval, with rates in between that declined linearly with time in development. This approach was meant to capture the essence of the risk-return staircase perspective expressed by Myers and others, and discussed below. The methodology described in Myers and Howe (1997) is actually quite different, but the OTA technique yielded results that would not be much different (for the same distribution of costs) than what one would have obtained with the correct methodology (Myers and Howe, 1997, p. 33).

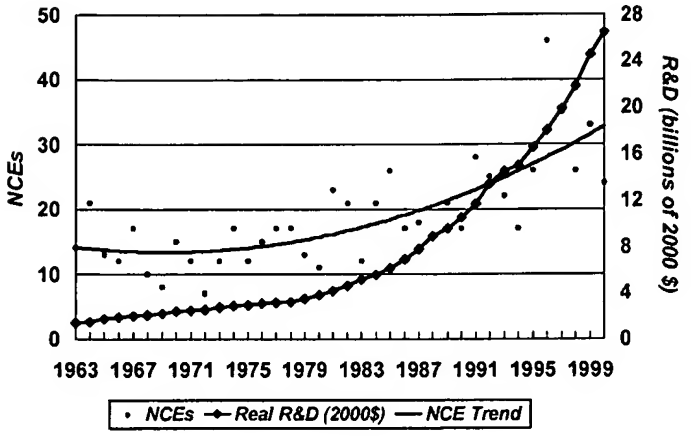


Fig. 1. Inflation-adjusted industry R&D expenditures (2000 dollars) and US new chemical entity (NCE) approvals from 1963 to 2000. Source of data: PhRMA (2001) and Tufts CSDD Approved NCE Database.

1.2. Aggregate data analyses

There have been no recent comprehensive studies of the cost of developing new pharmaceuticals from synthesis to marketing approval based on actual project-level data. However, aggregate data and data on parameters of the drug development process suggest that R&D costs have increased substantially since our earlier study. For example, the Pharmaceutical Research and Manufacturers of America (PhRMA, 2000) publishes an annual report on the R&D expenditures of its member firms that shows a continuous increase in outlays well in excess of inflation. Reports on specific components of the R&D process, such as the number of subjects in clinical trials (OTA, 1993; The Boston Consulting Group [BCG], 1993), also suggest an increase in the real cost of pharmaceutical innovation.

Published aggregate industry data suggest that R&D costs have been increasing. Fig. 1 shows reported aggregate annual domestic prescription drug R&D expenditures for members of the US pharmaceutical industry since 1963. The chart also shows the number of US new drug approvals by year. Given the much faster rate of growth of R&D expenditures, data such as these suggest that R&D costs have increased over time. However, they cannot be conclusive or precise. For one matter, the drug development process is known to be very lengthy. Thus, new drug approvals today are associated with R&D expenditures that were incurred many years prior. Ignoring the inherent lag structure underlying these data and simply dividing current R&D expenditures by the number of new drug approvals will in general yield inaccurate estimates.[2] Given a substantial increasing trend in R&D

[2] The estimates would also vary widely from year-to-year. For example, if we divided each year's real R&D expenditures by that year's number of NCE approvals, we would obtain US$ 1 billion for 2000, US$ 743 million for 1999, US$ 839 million for 1998, US$ 568 million for 1997, US$ 400 million for 1996, US$ 635 million for 1995, and US$ 878 million for 1994. While there is a general upward trend in such calculations, the year-to-year variability is not credible.

expenditures, such calculations will result in greatly exaggerated estimates of out-of-pocket cost per approval.

Secondly, even properly lagged time series would tend to be imprecise if aggregate industry data were used as reported. The industry data include expenditures on improvements to existing products. Thus, they would overestimate pre-approval development costs. On the other hand, they also do not incorporate all of the R&D on licensed-in drugs since firms or other organizations that are not members of the US trade association would have conducted some of the work. On that account the data would tend to underestimate costs. Therefore, R&D cost estimates based on project-level data are needed to assure a reasonable level of confidence in the accuracy of the results. We present results based on such data in this study.

The remainder of this paper is organized as follows. Section 2 describes the standard drug development paradigm, which serves as the structure through which the results are reported. Section 3 contains a description of the survey sample data and the population from which it was drawn. Section 4 describes the methodology used to derive R&D cost estimates. We present our base case pre-marketing approval R&D cost estimates in Section 5, as well as a comparison of our results with those of earlier studies to examine R&D cost trends. Section 6 provides sensitivity analyses for key parameters. Section 7 focuses on some extensions of the base case analyses: estimates of clinical development costs for approved drugs by therapeutic significance, estimates of post-approval R&D costs, and a tax analysis. Section 8 contains data and analyses that corroborate our results. Finally, we offer some conclusions in Section 9.

2. The new drug development process

New drug development can proceed along varied pathways for different compounds, but a development paradigm has been articulated that has long served well as a general model. The paradigm is explained in some detail elsewhere (DiMasi et al., 1991; US Food and Drug Administration [FDA], 1999). In outline form, the paradigm portrays new drug discovery and development as proceeding in a sequence of (possibly overlapping) phases. Discovery programs result in the synthesis of compounds that are tested in assays and animal models. It was not possible to disaggregate our data into discovery and preclinical development testing costs,[3] so for the purposes of this study discovery and preclinical development costs are grouped and referred to as preclinical costs.

Clinical (human) testing typically proceeds through three successive phases. In phase I, a small number of usually healthy volunteers[4] are tested to establish safe dosages and to gather information on the absorption, distribution, metabolic effects, excretion, and toxicity of the compound. To conduct clinical testing in the United States, a manufacturer must first

[3] The reported basic research expenditures by firm were highly variable, and suggest that different firms may categorize their pre-human testing expenditures somewhat differently. Thus, we report pre-human testing costs in one figure.

[4] In some therapeutic areas, testing is initially done on patients who have the disease or condition for which the compound is intended to be a treatment. This is ordinarily true in the cancer and AIDS areas.

file an investigational new drug application (IND) with the FDA. However, initiation of human testing can, and often does, occur first outside the United States.

Phase II trials are conducted with subjects who have the targeted disease or condition and are designed to obtain evidence on safety and preliminary data on efficacy. The number of subjects tested in this phase is larger than in phase I and may number in the hundreds. The final pre-approval clinical testing phase, phase III, typically consists of a number of large-scale (often multi-center) trials that are designed to firmly establish efficacy and to uncover side-effects that occur infrequently. The number of subjects in phase III trials for a compound can total in the thousands.

Once drug developers believe that they have enough evidence of safety and efficacy, they will compile the results of their testing in an application to regulatory authorities for marketing approval. In the United States, manufacturers submit a new drug application (NDA) or a biological license application (BLA) to the FDA for review and approval.

3. Data

Ten multinational pharmaceutical firms, including both foreign and US-owned firms, provided data through a confidential survey of their new drug R&D costs.[5] Data were collected on clinical phase costs for a randomly selected sample of the investigational drugs of the firms participating in the survey.[6] The sample was taken from a Tufts Center for the Study of Drug Development (CSDD) database of investigational compounds. Cost and time data were also collected for expenditures on the kind of animal testing that often occurs concurrently with clinical trials.[7] The compounds chosen were all self-originated; that is, their development up to initial regulatory marketing approval was conducted under the auspices of the surveyed firm.[8] Licensed-in compounds were excluded because non-survey firms would have conducted portions of the R&D.[9]

We also collected data from the cost survey participants on their aggregate annual pharmaceutical R&D expenditures for the period 1980–1999. The firms reported on total annual R&D expenditures broken down by expenditures on self-originated new drugs, on licensed-in or otherwise acquired new drugs, and on already-approved drugs. Annual expenditures on self-originated new drugs were further decomposed into expenditures during the pre-human and clinical periods.

The National Institutes of Health (NIH) support through their own labs and through grants to researchers in academic and other non-profit institutions a substantial amount of research

[5] Using pharmaceutical sales to measure firm size, four of the survey firms are top 10 companies, another four are among the next 10 largest firms, and the remaining two are outside the top 20 (PJB, 2000).

[6] A copy of the survey instrument is available upon request.

[7] Long-term teratogenicity and carcinogenicity testing may be conducted after the initiation of clinical trials.

[8] This does not preclude situations in which the firm sponsors trials that are conducted by or in collaboration with a government agency, an individual or group in academia, a non-profit institute, or another firm.

[9] Large pharmaceutical firms much more often license-in than license-out new drug candidates. Firms that license-in compounds for further development pay a price for that right through up-front fees, milestone payments, and royalty arrangements.

that expands fundamental knowledge about human biology (NIH, 2000; Scherer, 2000). This basic research sometimes results in leads that industrial researchers can capitalize on to assist them in discovering new therapeutic compounds.[10] Some new compounds investigated by pharmaceutical firms, however, originated in government or academic labs. It is unclear whether the discovery and early development costs for such compounds are similar to those for compounds originating in industrial labs. These drugs, though, represent a very small portion of the total number developed. For example, NIH (2000) found that of 47 FDA-approved drugs that had reached at least US$ 500 million in US sales in 1999, the government had direct or indirect use or ownership patent rights to only four of them.[11] In addition, we used a Tufts CSDD database supplemented by commercial databases to determine that of the 284 new drugs approved in the United States from 1990 to 1999,[12] 93.3% originated from industrial sources (either from the sponsoring firm or from another firm from which the compound was licensed or otherwise acquired). Government sources accounted for 3.2% of these approvals and academia and other non-profits accounted for the other 3.5%.[13]

The survey firms accounted for 42% of pharmaceutical industry R&D expenditures.[14] The survey compounds were selected at random from data contained in the Tufts CSDD database of investigational compounds for the firms that agreed to participate in the R&D cost survey. Of the 68 compounds chosen, 61 are small molecule chemical entities, four are recombinant proteins, two are monoclonal antibodies, and one is a vaccine. Initial human testing anywhere in the world for these compounds occurred during the period 1983–1994. Development costs were obtained through 2000.[15]

[10] The NIH also supports the development of research tools that drug developers find useful. In addition, it funds training for many scientists, some of whom eventually are employed in the industrial sector.

[11] The four drugs were developed in part through the use of NIH-funded patented technologies. Three of the four products are recombinant proteins, with two being the same drug produced by two different companies. Each of the relevant patented technologies was developed at academic or non-profit institutions with financial support from the NIH.

[12] The definition of a new drug used for this analysis is a therapeutic new molecular entity approved by the FDA's Center for Drug Evaluation and Research.

[13] The proportion of investigational drugs that derive from industrial sources is likely to be even higher, since acquired drugs have higher clinical approval success rates than do self-originated drugs (DiMasi, 2001b). Our cost survey firms were less reliant on licensing-in drugs from non-industrial sources than were firms as a whole; 98.8% of their new drug approvals during 1990–1999 were from industrial sources. DiMasi (2000) found markedly greater market entry of small niche pharmaceutical firms in the 1990s relative to earlier periods as measured by sponsorship of new chemical entity (NCE) approvals. A disproportionate share of the approvals obtained by these new entrants was for drugs that originated in academia.

[14] The data used were aggregate firm pharmaceutical R&D expenditures for the cost survey firms, as reported on our questionnaire, in comparison to PhRMA member firm R&D expenditures (1994–1997) on ethical pharmaceuticals, adjusted to global expenditure levels (PhRMA, 2001).

[15] Surveys were sent to 24 firms (some of whom have since merged). Twelve firms responded that they would participate in some form. The data that two firms ultimately provided were not useable. The 10 firms from which we used data provided information on 76 compounds. However, the data for eight of these compounds were not sufficiently comprehensive to use. The firms that did not participate in the survey cited a number of reasons for not doing so. The reasons included the extra demands that the transition effects of a relatively recent merger were placing on their relevant personnel, the time and expense of retrieving archival records in the manner required by the study, and difficulties in gathering the relevant data in a uniform manner because their accounting systems had changed significantly over the study period.

We selected a stratified random sample of investigational compounds. Stratification was based on the time elapsed since the origination of clinical trials and the current status of that testing. Reported costs were weighted to reflect the characteristics of the population, so that knowledge of the population from which the sample was drawn was needed. The population is composed of all investigational compounds in the Tufts CSDD investigational drug database that met study criteria: the compounds were self-originated and first tested in humans anywhere in the world from 1983 to 1994, and we had the information necessary to classify them according our strata. We found 538 investigational drugs that met these criteria. Of these compounds, 82 (15.2%) have been approved for marketing, 9 (1.7%) had NDAs or BLAs that were submitted and are still active, 5 (0.9%) had NDAs or BLAs submitted but abandoned, 227 (42.2%) were terminated in 4 years or less from the initiation of clinical trials, 172 (32.0%) were terminated more than 4 years after the start of clinical testing, and 43 (8.0%) were still in active testing as of the most recent check (31 March 2001).

Some firms were not able to provide full phase cost data for every new drug sampled. For example, phase I cost data were available for 66 of the 68 new drugs. However, we had some phase cost data for every drug in the sample. In addition, five compounds were still active at the time of the study. For these drugs it is possible that there will be some future costs for the drug's most recent phase. Thus, for this reason our cost estimates may be somewhat conservative. However, given the small number of drugs in this category and the fact that the impact would be on only one phase for each of these drugs, our overall cost estimates are not likely to be materially affected.

4. Methodology for estimating new drug development costs

The approach that we use to estimate development costs is similar to that described in our earlier work (DiMasi et al., 1991). We will outline here the general methodology for developing an overall cost estimate. In describing the approach, it will be clear that cost estimates for important components of the drug development process will also be derived along the way.

The survey sample was stratified to reduce sampling error. Results from previous analyses suggested that the variability of drug costs tends to increase with the development phase or the amount of time that a drug spends in testing (Hansen, 1979; DiMasi et al., 1991). Costs for successful drugs (i.e. those that achieve regulatory approval) also tend to be higher and more variable than those for drug failures. Thus, we based our strata on the length of time that failed compounds were in clinical testing and whether or not a compound had reached the stage in which an application for marketing approval had been filed with the FDA.[16]

[16] Specifically, we used four strata: compounds that failed in 4 years or less of clinical testing; compounds that failed after more than 4 years had elapsed from initial human testing; compounds for which an NDA or a BLA had been submitted to the FDA; and compounds that were still in active testing (as of 30 March 2001). Compounds for which an application for marketing approval had been submitted or which had been abandoned after lengthy testing were deliberately oversampled. The reported sample values were then weighted, where the weights were determined so that the sample perfectly reflects the population in terms of the four strata.

4.1. Expected costs in the clinical period

Since new drug development is a risky process, with many compounds failing for every one that succeeds, it is necessary to analyze costs in expected value terms. The total clinical period cost for an individual drug can be viewed as the realization of a random variable, c. Given that it is not certain that development of a randomly selected investigational compound will proceed to a given phase, we may define expected clinical costs for a randomly selected investigational drug to be $C = E(c) = p_{\mathrm{I}}\mu_{\mathrm{I}|\mathrm{e}} + p_{\mathrm{II}}\mu_{\mathrm{II}|\mathrm{e}} + p_{\mathrm{III}}\mu_{\mathrm{III}|\mathrm{e}} + p_{\mathrm{A}}\mu_{\mathrm{A}|\mathrm{e}}$, where p_{I}, p_{II}, and p_{III}, are the probabilities that a randomly selected investigational compound will enter phases I–III, respectively, p_{A} the probability that long-term animal testing will be conducted during the clinical trial period, and the μ's are conditional expectations. Specifically, $\mu_{\mathrm{I}|\mathrm{e}}$, $\mu_{\mathrm{II}|\mathrm{e}}$, $\mu_{\mathrm{III}|\mathrm{e}}$, and $\mu_{\mathrm{A}|\mathrm{e}}$ are the population mean costs for drugs that enter phases I–III, and clinical period long-term animal testing, respectively.

Weighted mean phase costs derived from the cost survey data were used to estimate the conditional expectations. A description of how the probabilities were estimated is presented in the next section. Assuming that the estimated mean phase costs and success probabilities are stochastically independent, the estimated expected value is an unbiased estimate of the population expected value.

4.2. Clinical success and phase attrition rates

An overall clinical approval success rate is the probability that a compound that enters the clinical testing pipeline will eventually be approved for marketing. Attrition rates describe the rate at which investigational drugs fall out of testing in the various clinical phases. A phase success rate is the probability that a drug will attain marketing approval if it enters the given phase. A phase transition probability is the likelihood that an investigational drug will proceed in testing from one phase to the next. All of these probabilities can be estimated from data in the Tufts CSDD database of investigational drugs from which our survey sample was drawn.

The clinical approval success rate was estimated using a two-stage statistical estimation process that has been described in detail elsewhere (DiMasi et al., 1991; DiMasi, 2001b). The data used here consist of the investigational drugs in the Tufts CSDD database that were first tested in humans anywhere in the world from 1983 to 1994, with information on their status (approval or research abandonment) obtained through early 2001. Given that some of these investigational drugs were still in active testing at the end of the study period, some of the data are right-censored. Survival analysis can be applied in such a situation, where survival indicates that a drug has not reached its ultimate fate (either approval or abandonment).

The Tufts CSDD database of investigational compounds contains information on the latest phase that an abandoned compound was in when it was terminated. These data were used to determine the distribution of research terminations by phases.[17] These results,

[17] A small proportion of the compounds in the database were either still in clinical development (8.0%) or had an NDA or BLA filed but not yet approved (1.7%). For those drugs in these groups that will eventually fail, their abandonment will tend to occur in later testing phases. To deal with the potential bias in the estimated distribution of research terminations that would result from using just those compounds that had been abandoned by the end of

together with the estimated overall clinical approval success rate were used to provide estimates of the probability that an investigational drug will enter a given phase, phase attrition rates, and phase transition probabilities. The estimated overall clinical approval success rate and the probabilities of entering various phases provide results with which estimates can be derived that include the cost of drugs that fail to make it through the development process. Specifically, we use the probabilities of entering a phase to estimate the expected out-of-pocket clinical cost per investigational drug. Adding the out-of-pocket preclinical cost estimate described below yields an estimate of total out-of-pocket cost per investigational drug. Dividing this estimate by the overall clinical success rate yields our estimate of out-of-pocket cost per approved drug.

4.3. Out-of-pocket discovery and preclinical development costs

Many costs incurred prior to clinical testing cannot be attributed to specific compounds. Thus, aggregate level data at the firm level were used to impute costs per drug for R&D incurred prior to human testing. Specifically, time series data for each surveyed firm on spending on pre-human R&D and on human testing for 1980–1999 were obtained, and a ratio of pre-human R&D expenditures to human testing expenditures was determined based on an appropriate lag structure (on average, pre-human R&D expenditures should occur years prior to the associated human testing costs). This ratio was then multiplied by an estimate of out-of-pocket clinical cost per drug, which is based on the project-level data, to yield an estimate of the pre-human R&D cost per new drug.[18]

4.4. Capitalized costs: development times and the cost-of-capital

Given that drug development is a very lengthy process, the full cost of drug development should depend significantly on the timing of investment and returns. Full cost estimates require a capitalization of the stream of out-of-pocket costs to some point (the date of marketing approval is the standard). To do so, one needs a timeline for a representative drug. The timeline is constructed from information on average phase lengths and the average gaps and overlaps between successive phases in a Tufts CSDD database of approved new drugs and in our cost survey. The periods considered are the time from synthesis to human testing,

---

the study period, we statistically predicted whether each open compound (still in clinical testing) would eventually fail. To do so, we evaluated an estimated conditional approval probability function (probit specification) at the number of years that the compound had been in testing. Failures were taken to occur in the latest reported testing phase. Summing the failure probabilities by phase gives us additional terminations by phase. The distribution of research terminations by phase was adjusted accordingly. Compounds that had reached the NDA/BLA phase likely have a very high probability of success. DiMasi (2001a) found very high approval rates for NDA submissions, with an increasing trend. To be conservative, we assumed that all of the compounds with still active NDAs or BLAs would be approved. This leads to lower cost estimates than would be the case if the same procedure for determining failure that was used for compounds still in testing had been used instead. However, given the very small number of compounds in the active NDA/BLA category, the impact on the results is trivial.

[18] The survey firms were asked to indicate whether charges for corporate overhead unrelated to R&D appear in their R&D budget data, and, if so, to estimate what share of expenditures they represent. Two firms reported that they did, and so we reduced the aggregate and project-level data for those firms according to their reported shares for corporate overhead.

the three clinical phases, an animal testing phase concurrent with clinical development, and the length of time from submission of an NDA/BLA to NDA/BLA approval.

Whereas the survey data cover a development period that yielded approvals from 1990 to 2001, the bulk of the approvals occurred in the mid to late 1990s. Thus, we estimated phase lengths, gaps, and overlaps for self-originated new drugs that were approved during 1992–1999. The data included therapeutic biopharmaceuticals, as well as small molecule drugs.[19] Once a timeline is established and out-of-pocket costs are allocated over that timeline, the expenditures must be capitalized at an appropriate discount rate. The discount rate should be the expected return that investors forego during development when they invest in pharmaceutical R&D instead of an equally risky portfolio of financial securities. Empirically, such a discount rate can be determined by examining stock market returns and debt-equity ratios for a representative sample of pharmaceutical firms over a relevant period. The resulting discount rate is an average company cost-of-capital. We describe the estimation of our base case cost-of-capital in Section 5.2 below.

We assume that phase costs are distributed uniformly over the phase length and apply continuous compounding to the point of marketing approval. Summing these capitalized preclinical and clinical capitalized cost estimates yields a total capitalized cost per investigational drug. Dividing by the overall clinical success rate results in our estimate of the total capitalized cost per approved new drug. This estimate is a measure of the full resource cost needed, on average, for industry to discover and develop a new drug to the point of marketing approval.

5. Base case R&D cost estimates

5.1. Out-of-pocket clinical cost per investigational drug

Given the method of weighting reported costs as described in Section 4, weighted means, medians, and standard deviations were calculated and are presented in Table 1.[20] Mean

[19] The percentage of all self-originated new compound approvals that are for biopharmaceuticals is substantially larger than is the proportion of either self-originated approvals or investigational compounds that are for biopharmaceuticals in the Tufts CSDD investigational drug database. The survey firms in this database are predominantly traditional pharmaceutical firms. Thus, we estimate clinical phase lengths and approval phase times for new chemical entities and biopharmaceuticals separately and compute a weighted average of the mean phase lengths, where the weights are the shares of self-originated investigational compounds in the Tufts CSDD database for each of these compound types.

[20] For five of the sample drugs, the survey firms were not able to disaggregate costs for two successive clinical phases (i.e. either phases I and II or phases II and III). We developed a two-stage iterative process for imputing phase costs for these drugs. To illustrate, suppose that the firm combined phases II and III costs for a specific drug. For a year during which the drug was in both phase II and III testing, let m_{II} = number of months the drug was in phase II only, m_{III} = number of months the drug was in phase III only, m_0 = number of months the drug was in both phases, T = total clinical phase cost for the drug during the year, and cr = ratio of weighted monthly phase III to phase II cost for drugs where phase costs were disaggregated. Imputed phase II cost, x_{II}, can then be defined as $x_{II} = (m_{II} + \text{cr} \cdot m_0)T/(m_{II} + \text{cr} \cdot m_{III} + [1 + \text{cr}] \cdot m_0)$. Imputed phase III cost is determined as $x_{III} = \text{cr} \cdot x_{II}$. The same approach was used when phase I and II costs were combined by the responding firm. To further refine the results, we included the imputed costs for the five drugs from the first stage and recomputed the phase cost ratios to determine second stage values for the imputed costs. The results for imputed costs barely changed between the first and the second iterations.

Table 1
Average out-of-pocket clinical period costs for investigational compounds (in millions of 2000 dollars)[a]

Testing phase	Mean cost	Median cost	Standard deviation	N[b]	Probability of entering phase (%)	Expected cost
Phase I	15.2	13.9	12.8	66	100.0	15.2
Phase II	23.5	17.0	22.1	53	71.0	16.7
Phase III	86.3	62.0	60.6	33	31.4	27.1
Long-term animal	5.2	3.1	4.8	20	31.4	1.6
Total						60.6

[a] All costs were deflated using the GDP Implicit Price Deflator. Weighted values were used in calculating means, medians, and standard deviations.

[b] *N*: number of compounds with full cost data for the phase.

cost per investigational drug entering a phase increases substantially by clinical phase, particularly for phase III, which is typically characterized by large-scale trials. In comparison to the previous study (DiMasi et al., 1991), mean phase I cost is moderately higher relative to the other phases. While the ratio of mean phase III cost to mean phase I cost was 6.0 for the previous study, it was 5.7 here. Similarly, the ratio of mean phase II to phase I cost was 1.9 for the earlier study, but was 1.5 for this study. The higher relative phase I cost is consistent with other data that indicate that the growth in the number of procedures per patient was much greater for phase I than for the other phases during the 1990s.[21]

Mean clinical phase costs increased approximately five-fold in real terms between the studies. However, in comparison, long-term animal testing costs incurred during the clinical period increased by only 60%. Thus, increases in out-of-pocket clinical period costs were driven heavily by increases in human trial, as opposed to animal testing, costs. This suggests that preclinical animal studies may also have not increased at anywhere near the same rate as have clinical trial costs. The results also indicate that development costs have become more uniform across drugs.

This is indicated by two comparisons with the results from the previous study. The ratio of mean to median phase cost decreased 50% for phase I, 22% for phase II, and 13% for phase III for the present study in comparison to the earlier study. Thus, the data are less skewed. The coefficients of variation for the phases also declined. They are 60% lower for phase I, 29% lower for phase II, and 36% lower for phase III.

Estimates of the probability that an investigational drug will enter a phase were obtained from statistical analysis of information in the Tufts CSDD database of investigational compounds for drugs that met study criteria. They are shown in Table 1 and are used to obtain the expected phase costs in the last column. The probabilities are lower in comparison to the previous study (75.0% for phase II, 36.3% for phase III, and 56.1% for long-term animal testing). Lower probabilities of entering a phase will, other things being equal, result in lower expected costs. Thus, while the mean phase costs for drugs entering a phase are

[21] One of the authors obtained data from DataEdge, LLC on the number of procedures administered to patients by phase from 1990 to 1997. The data were based on information in the clinical trial grants of a very large number of pharmaceutical firms. During this period, the number of procedures per patient increased 27% for phase III, 90% for phase II, and 120% for phase I.

approximately five times higher in this study, the expected cost per investigational drug is only four times higher.

Alternative probability estimates for the same data make clear how reductions in drug development risks hold down development costs. Our earlier study showed proportionately fewer failures in phase I (32.5% versus 37.0%) and proportionately more failures in phase III (17.1% versus 12.6%); the share for phase II was identical. Thus, given a similar overall clinical success rate, the evidence suggests that over time firms became better able to weed out failures (clinical or economic) early in the process. A similar scenario holds when we examine phase transition probabilities. In the earlier study, a larger percentage of investigational drugs made it to phase II (75.0% versus 71.0%) and a smaller percentage proceeded from phase III to marketing approval (63.5% versus 68.5%).

5.2. *Cost-of-capital estimates*

In our earlier paper (DiMasi et al., 1991), we employed a 9% real cost-of-capital based on a capital asset pricing model (CAPM) analysis for a representative group of pharmaceutical firms during the 1970s and early 1980s. A real rather than a nominal cost-of-capital is appropriate in our analysis since R&D costs are expressed in constant 2000 dollars. The real cost-of-capital in pharmaceuticals has increased since the mid-1980s primarily as a result of higher real rates of return required by holders of equity capital during the 1990s.

In the present analysis, we compute a weighted cost-of-capital for each firm in a representative group of pharmaceutical firms for the 1980s and 1990s, where the weights are based on the firm's market value of debt and equity. For most major pharmaceutical firms, debt securities account for less than 10% of market valuation, so that the equity cost-of-capital component is the dominant element of the weighted cost-of-capital for this industry. At the request of the OTA, Myers and Shyam-Sunder (1996) estimated the cost-of-capital for the pharmaceutical industry during the 1970s and 1980s using a standard CAPM approach. Their methodology is the basis for our updated analysis.

In our R&D cost analysis we have a sample of new drugs that began clinical trials in the mid-1980s through the early 1990s, and which have an average market introduction point in the late 1990s. Hence a relevant time period for our cost-of-capital measure is 1985–2000. Accordingly, we estimated the cost-of-capital at roughly 5-year intervals beginning in January 1985 and ending in January 2000. The results of our analysis are summarized in Table 2.

The nominal cost-of-capital in 1985 and 1990 are based on Myers and Shyam-Sunder's analysis for the OTA. The 1994 value is from Myers and Howe (1997). The 2000 nominal cost-of-capital (COC) value is based on our own estimation, employing a sample of firm and data sources comparable to those used in the prior work of Myers and colleagues. As can be seen in Table 2, the nominal cost-of-capital for pharmaceutical firms has remained relatively stable in this period in the range of 14–16%, with a mean of approximately 15%.[22]

[22] We undertook an informal survey of major pharmaceutical firms in mid-2001 with respect to the hurdle rate that they used in their R&D investment decisions. This survey of six firms yielded (nominal) hurdle rates from 13.5 to over 20%. This indicates that a 15% nominal COC rate is within the range of hurdle rates utilized by major pharmaceutical firms for their actual R&D investments.

Table 2
Nominal and real cost-of-capital (COC) for the pharmaceutical industry, 1985–2000

	1985	1990	1994	2000
Nominal COC (%)[a]	16.1	15.1	14.2	15.0
Inflation rate (%)[b]	5.4	4.5	3.1	3.1
Real COC (%)	10.8	10.6	11.1	11.9

[a] The nominal values for 1985 and 1990 are based on Myers and Shyam-Sunder (1996). The nominal value for 1994 is taken from Myers and Howe (1997). The 2000 nominal value is based on our own computations using comparable samples and data sources.

[b] The inflation rate for 1985 is taken from Myers and Howe (1997), the rate for 1990 is a 5-year average centered on January 1990 and is based on the CPI-U, the rate for 1994 and 2000 is the long-term inflation rate from 1926 to 2000 (Ibbotson Associates, 2001, p. 17).

To obtain a real cost-of-capital, we subtracted the expected rate of inflation from the nominal cost-of-capital. For this purpose, Myers and Shyam-Sunder (1996) used the expected rate of inflation from a special consumer survey performed in the 1980s. We also used this value in Table 2 for the 1985 period. For 1990 we utilized a 5-year moving average of actual inflation rates centered around the year in question to estimate expected rates of inflation. For 1994 and 2000 we used the long-term inflation rate (1926–2000) in Ibbotson and Associates (2001) of 3.1% to compute the values in Table 2.[23]

The real cost-of-capital for the pharmaceutical industry over this period, using the CAPM model, varies from 10.6 to 12.0%. The mean cost-of-capital in this period was just over 11%. Hence, 11% is the baseline value that we employed in our R&D cost estimates.[24] However, as in prior studies, we did sensitivity analysis around this value in order to determine how our baseline R&D cost estimates are affected by changes in the cost-of-capital.

5.3. *Capitalized clinical cost per investigational drug*

To calculate opportunity cost for clinical period expenditures we estimated average phase lengths and average gaps or overlaps between successive clinical phases. Mean phase lengths and mean times between successive phases are shown in Table 3. The time between the start of clinical testing and submission of an NDA or BLA with the FDA was estimated to be 72.1 months, which is 3.5 months longer than the same period estimated in the previous study. However, the time from the start of clinical testing to marketing approval in our timeline for a representative drug averaged 90.3 months for the current study, compared to

[23] Inflation rates were particularly low in the 1990s, and 5-year moving averages were below the long-term rate. Since the 1990s represented a marked change in the inflation rate from earlier decades, and inflationary expectations may not adjust immediately to the new experience, we used the long-term inflation rate rather than 5-year moving averages for this period.

[24] This yields conservative estimates of the cost of capital from several perspectives. One important point concerns the fact that many major pharmaceutical firms have large positive cash balances and are actually net lenders rather than net borrowers (i.e. they have a negative debt ratio). Incorporating this point into their CAPM analysis for January, 1990, causes the estimated nominal value of the cost of capital to increase by almost a full percentage point (see Myers and Shyam-Sunder, 1996, p. 223). In addition, as noted in footnote 4, many firms appear to use higher costs of capital in their R&D investment decisions than what emerges from this CAPM analysis.

Table 3
Average phase times and clinical period capitalized costs for investigational compounds (in millions of 2000 dollars)[a]

Testing phase	Mean phase length	Mean time to next phase	Capitalized mean phase cost[b,c]	Capitalized expected phase cost[b,c]
Phase I	21.6	12.3	30.5	30.5
Phase II	25.7	26.0	41.6	29.5
Phase III	30.5	33.8	119.2	37.4
Long-term animal	36.5	–	9.5	3.0
Total				100.4

[a] All costs were deflated using the GDP Implicit Price Deflator. Weighted values were used in calculating means, medians, and standard deviations for costs and phase times. Phase times are given in months.

[b] The NDA approval phase was estimated to be 18.2 months. Animal testing was estimated to begin 4.2 months after the initiation of phase I.

[c] Costs were capitalized at an 11% real discount rate.

98.9 months for the earlier study. The difference is accounted for by the much shorter FDA approval times in the mid to late 1990s that were associated with the implementation of the *Prescription Drug Use Fee Act of 1992*. While the approval phase averaged 30.3 months for the earlier paper's study period, that phase averaged only 18.2 months for drugs covered by the current study.

Other things being equal, the observed shorter times from clinical testing to approval yield lower capitalized costs relative to out-of-pocket costs. However, the discount rate that we used for the current study is also higher than for the previous study (11% versus 9%). The two effects work in offsetting ways. On net, there was very little difference between the studies in the ratio of mean capitalized to out-of-pocket cost for the individual clinical phases.[25]

5.4. Clinical cost per approved new drug

Although average cost estimates for investigational drugs are interesting in their own right, we are mainly interested in developing estimates of cost per approved new drug. To do so, we need an overall clinical approval success rate. Our statistical analysis of compounds in the Tufts CSDD database of investigational drugs that met study criteria yielded a predicted final clinical success rate of 21.5%. Applying this success rate to our estimates of out-of-pocket and capitalized costs per investigational drug results in estimates of cost per approved new drug that link the cost of drug failures to the successes.

Aggregating across phases, we find that the out-of-pocket clinical period cost per approved new drug is US$ 282 million and the capitalized clinical period cost per approved new drug is US$ 467 million. These costs are more than four-fold higher than those we found in our previous study.

[25] The ratios of capitalized to out-of-pocket cost for the earlier study were 1.9, 1.7, 1.4, and 1.6 for phases I–III, and animal testing, respectively. For this study, we found the ratios to be 2.0, 1.8, 1.3, and 1.8 for phases I–III, and animal testing, respectively.

5.5. Preclinical out-of-pocket and capitalized costs per approved drug

The preclinical period, as defined here, includes discovery research as well as preclinical development. As noted above, not all costs during this period can be allocated to specific compounds. To deal with this issue, we analyzed aggregate annual firm expenditures on self-originated new drugs by the preclinical and clinical periods. We gathered data on aggregate expenditures for these periods from survey firms for 1980–1999. Both times series tended to increase over time in real terms. Given this outcome, and the fact that the clinical expenditures in 1 year will be associated with preclinical expenditures that occurred years earlier, the ratio of total preclinical expenditures to total R&D (preclinical plus clinical) expenditures over the study period will yield an overestimate of the share of total cost per new drug that is accounted for by the preclinical period. To accurately estimate this share we built in a lag structure that associates preclinical expenditures with clinical expenditures incurred some time later. Using data in the Tufts CSDD database of approved drugs, we estimated the average time from synthesis of a compound to initial human testing for self-originated drugs to be 52.0 months. Our analysis of clinical phase lengths and phase gaps and overlaps indicates a period of 68.8 months over which clinical period development costs are incurred. We approximate the lag between preclinical and clinical expenditures for a representative new drug as the time between the midpoints of each period. This yields a lag of 60.4 months, or approximately 5 years. Thus, we used a 5-year lag in analyzing the aggregate expenditure data. Doing so resulted in a preclinical to total R&D expenditure ratio of 30%. This share was applied to our clinical cost estimates to determine corresponding preclinical cost estimates. Given the estimates of out-of-pocket and capitalized clinical cost per approved new drug noted in Section 5.4, we can infer preclinical out-of-pocket and capitalized costs per approved new drug of US$ 121 and 335 million, respectively. The results are very robust to different values for the length of the lag structure. For example, if we assume a lag of 4 years instead of 5 years, then out-of-pocket preclinical costs would be 9.8% higher. Alternatively, if we assume a 6-year lag, then out-of-pocket preclinical costs would be 9.3% lower.

5.6. Total capitalized cost per approved drug

Our full cost estimate is the sum of our preclinical and clinical period cost estimates. Our base case out-of-pocket cost per approved new drug is US$ 403 million, while our fully capitalized total cost estimate is US$ 802 million. Time costs, thus, account for 50% of total cost. This share is nearly identical to one that we found in our previous study (51%). This is the case even though the time cost shares for both the clinical and preclinical periods are somewhat higher for this study. The explanation for this seeming inconsistency is that time costs are relatively greater for preclinical expenditures since they are incurred earlier in the process, but the preclinical share of total costs is lower for the present study.

5.7. Trends in R&D costs

Fig. 2 presents the primary results (capitalized preclinical, clinical, and total cost per approved new drug) for the previous two studies and for our current study. In inflation-adjusted

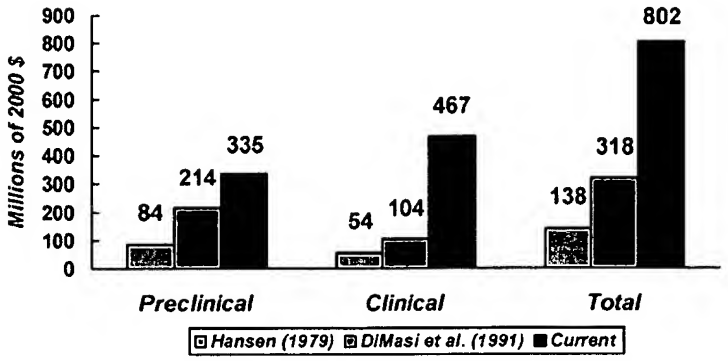


Fig. 2. Trends in capitalized preclinical, clinical and total cost per approved new drug.

terms, total capitalized cost was 2.3 times higher for the previous study in comparison to the first study. Real total capitalized cost per approved new drug for the current study is 2.5 times higher than for the previous study. However, the samples for these studies include drugs that entered clinical testing over periods that are not uniformly dispersed. In addition, while the samples were chosen on the basis of when drugs entered clinical testing, changes over time in the average length of the development process make ascribing differences in the study periods according to the year of first human testing problematic. An alternative is to determine an average approval date for drugs in each study's sample and use the differences in these dates to define the time differences between the studies. This will allow us to determine annual cost growth rates between successive studies.

Drugs in the current study sample obtained FDA marketing approval from 1990 to 2001, with the vast majority of the approvals occurring between 1992 and 2000. The mean and median approval date for drugs in the current study's sample was in early 1997. For the previous study, we reported that the average approval date was in early 1984. Thus, we used 13 years as the relevant time span between the studies and calculated compound annual rates of growth between the two studies accordingly.

Hansen (1979) did not report an average approval date; however, we can infer a period difference by noting the sample selection criteria and the difference in development times between that study and the DiMasi et al. (1991) study. The sample selection criteria for DiMasi et al. (1991) involved a 7-year shift in initial clinical testing relative to Hansen (1979). However, the estimated time from the start of clinical testing to marketing approval was 2.3 years longer for the DiMasi et al. (1991) study. Thus, we use 9.3 years as the difference between the study periods for these two studies.

Using these period differences, we found that the compound annual growth rates in total out-of-pocket cost per approved drug were quite similar across the studies (Table 4). The growth in total costs, however, masks substantial differences in growth rates for the preclinical and clinical periods. While out-of-pocket preclinical expenditures continued to grow in real terms, its growth rate for the current study relative to the previous one declined by two-thirds in comparison to the growth rate for the first two studies. Conversely, the growth rate for clinical period expenditures approximately doubled for the two most recent studies.

Table 4
Compound annual growth rates in out-of-pocket and capitalized inflation-adjusted costs per approved new drug[a]

Period	Out-of-pocket			Capitalized		
	Preclinical (%)	Clinical (%)	Total (%)	Preclinical (%)	Clinical (%)	Total (%)
1970–1980	7.8	6.1	7.0	10.6	7.3	9.4
1980–1990	2.3	11.8	7.6	3.5	12.2	7.4

[a] Costs for the 1970s approvals are from Hansen (1979), costs for the 1980s approvals are from DiMasi et al. (1991), and costs for the 1990s approvals are from the current study.

Annual growth rates for capitalized costs are also shown in Table 4. The results show a substantially higher growth rate for clinical costs for the two most recent analyses. However, while the growth rate for total out-of-pocket cost per approved drug was slightly greater for the two most recent studies, the growth rate in total capitalized cost was two percentage points higher between the first and second study than between the second and third. This is so, despite the fact that the discount rate increased one percentage point between the first two studies, but two percentage points between the last two. The growth rate in capitalized costs, however, is driven more by the fact that preclinical costs have a lower share of total out-of-pocket costs in the current study than in the previous studies, and time costs are necessarily proportionately more important for preclinical than for clinical expenditures.

6. Sensitivity analysis

6.1. *Effects of parameter changes*

We undertook sensitivity analyses for several of the key parameters that underlie the cost estimates. Fig. 3 shows how preclinical, clinical, and total capitalized costs would vary by discount rate at half-percentage point intervals. The values for a zero percent discount rate are out-of-pocket costs. In the neighborhood of our base case discount rate (11%), clinical cost changes by about US$ 10 million, preclinical cost changes by about US$ 15 million,

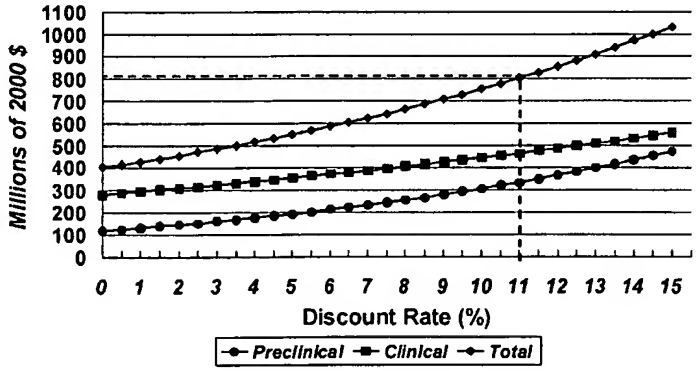


Fig. 3. Capitalized preclinical, clinical, and total costs per approved new drug by discount rate.

and total cost changes by about US$ 25 million for every half of one percent shift in the discount rate. In our previous study, the base case discount rate was 9%. At a 9% discount rate, total capitalized cost here is US$ 707 million, or 11.8% less than our base case result. The results in section 5.3 provide some support for an even higher discount rate than our base case value. At a 12% discount rate, total capitalized cost per approved new drug is US$ 855 million, or 6.6% higher than our base case result.

The clinical approval success rate is another key parameter. We analyzed the effects of an approximate 10% change in the success rate at various discount rates. A higher success rate has a somewhat smaller impact on total cost than does a correspondingly lower success rate. At our base case discount rate, total capitalized cost for a success rate of 23.5% is US$ 734 million, or 8.5% lower than our base case result. At a success rate of 19.5%, total capitalized cost is US$ 885 million, or 10.3% higher than our base case result. The estimated clinical success rate for our previous study was 23.0%. At that success rate, total capitalized cost here is US$ 750 million, or 6.5% less than our base case result.[26]

The methodology for determining the total capitalized cost estimate is dependent on values for 20 parameters. However, not all of them are independent of one another. It is possible to determine total capitalized cost from estimates of 16 parameters. To get a measure of statistical error for overall cost, we performed a Monte Carlo simulation (1000 trials) for total capitalized cost by taking random draws from the sampling distributions of the 16 parameters and computing a total cost estimate for each simulation trial.[27] Ninety-five percent of the total cost estimates for the simulation fell between US$ 684 and 936 million, 90% fell between US$ 705 and 917 million, and 80% fell between US$ 717 and 903 million.[28] The interquartile range was US$ 757–854 million.

[26] These analyses indicate what the results would be if the clinical success rate is changed, while other parameters remain the same. If the phase attrition rates are adjusted to be consistent with the new clinical success rate while maintaining the same distribution of failures across phases, then the differences in cost are somewhat lower. For example, if the clinical success rate is 23.0% and phase attrition rates are altered accordingly, total capitalized cost is 5.6% lower (5.1% lower if account is also taken of estimated differences in phase costs between the failures and successes in the sample [see the following section]).

[27] The clinical success rate parameter is determined from the values of four asymptotically normal coefficient estimates. We performed an initial Monte Carlo simulation for the clinical success rate using these coefficient estimates and their standard errors to obtain a sampling distribution for the success rate. The sampling distribution for the discount rate was chosen by assumption. Given that the base case choice of discount rate may be somewhat conservative (see the discussion above), we chose a triangular distribution for the discount rate that varied from 10.0 to 12.5%, with the modal value for the distribution chosen so that the mean discount rate is approximately 11.0% in the simulations for total capitalized cost. The other sampling distributions were for estimated means and binomial probabilities. Finite population correction factors were applied to the standard errors.

[28] The simulation was conducted assuming statistical independence for the parameters. The out-of- pocket phase cost, development time, and success and attrition rate parameters were estimated from separate datasets, and so their independence of one another is likely. It is possible that out-of- pocket phase costs are correlated. We therefore also conducted a simulation using the estimated correlations across phases for those pairs that were found to have correlations that were statistically significantly different from zero (phases I and II [0.496], phases II and III [0.430], phase II and long-term animal testing [0.656]). This increased the variability of the total capitalized cost estimates only slightly. Specifically, the coefficient of variation increased from 0.088 to 0.099. The main simulation results were affected most by variability in individual phase costs, and least by variability in development times. The coefficient of variation when only development times vary, when only the discount rate varies, when only success and attrition rates vary, and when only out-of-pocket phase costs vary were 0.015, 0.035, 0.044, and 0.065, respectively.

6.2. Variable discount rates

Myers and others (Myers and Shyam-Sunder, 1996; Myers and Howe, 1997) have argued that the cost-of-capital for R&D should decline over the development process as a step function. They termed the relationship a risk-return staircase. In the case of pharmaceutical R&D, the staircase is not related to the usual notions of risk in pharmaceutical development (i.e. the probabilities of approval at different points in the process). These technical risks can be diversified away by investors, who can spread their investments over many firms. Rather, the rationale has to do with the notion that at any point in the development process future R&D costs serve as a kind of leverage, or debt, if the firm wishes to proceed with development and market a product. A more levered position amplifies risk and is associated with a higher cost-of-capital for investors. Since the leveraging declines over the development process, so does the cost-of-capital. Technical risks play a role only in that they affect expected future costs.

The valuation problem may also be viewed as a compound option pricing problem. The firm effectively faces call options at decision points during development, where the exercise price is the cost of future R&D. Myers and Howe (1997) suggest a means for dealing with the problem that reduces the informational requirements to knowledge of two-discount rates. One of these is the discount rate for net revenues on a marketed drug (r_{NR}). The other is the discount rate on future costs (r_{FC}). The rate for net revenues should be somewhat less than the overall company COC. The rate for future costs, being an expected return on what is nearly a fixed debt obligation, is likely lower. Under certain assumptions, the Myers and Howe (1997) two-discount rate method yields the same results as the more complex compound options valuation. We view this approach to discounting as experimental for our purposes. To our knowledge, no pharmaceutical firm uses such an approach for its project evaluations. In addition, although they may be guided by real world information, the selections of the two-discount rates are judgment calls.[29]

For purposes of comparison, we did compute drug R&D costs with the Myers and Howe (1997) two-discount rate method. Their base case values for r_{NR} (9%) and r_{FC} (6%) were meant to be relevant for 1994, which corresponds roughly with the middle of our study period. Thus, we computed the total capitalized pre-approval cost per approved drug using these values and other close combinations in a sensitivity analysis. At the Myers and Howe (1997) base case values, total capitalized cost is marginally higher than our estimate computed at an 11% COC (US$ 815 million). However, the total capitalized cost estimate is US$ 955 million when a 10% discount rate is used for r_{NR} and a 5% discount rate is used for r_{FC}. Conversely, at an 8% discount rate for r_{NR} and a 7% discount rate for r_{FC}, the total cost estimate is US$ 696 million.

[29] For their financial life-cycle simulation model, Myers and Howe (1997) chose base case values for r_{NR} and r_{FC} partly on the basis of judgment and partly because these values generated realistic company costs-of-capital for mature pharmaceutical firms in their simulations. These simulations required assumptions about revenue distributions and other factors that affect profitability.

7. Extensions to the base case

The base case results on overall pre-approval drug development costs can be extended in several interesting ways. Our base case results link the costs of the failures to the successes. We can provide estimates of the clinical period cost of taking a successful drug all the way to approval by examining the data for the approved drugs in the sample. This also allows us to obtain some evidence on costs for the more medically significant products (according to what is known at the time of approval) by using an FDA prioritization ranking for approved drugs. We can also use data collected from our survey to estimate R&D expenditures on new drugs subsequent to original marketing approval. Finally, we can examine what impact tax policies and procedures have had on the effective cost of pharmaceutical R&D for pharmaceutical firms.

7.1. Development costs for successes

As our results indicate, development costs vary across drugs. Thus, it is worthwhile to examine specific subclasses of drugs, where one may reasonably conjecture that the cost structure is different than it is for drugs as a whole. In particular, we investigated the clinical cost structure for successful drugs (i.e. drugs that have made it through testing and obtained marketing approval from the FDA). We also examined these data classified by an FDA rating of therapeutic significance for drug approvals.

Of the 68 drugs in our sample, 27 have been approved for marketing. We had complete phase cost data for 24 of the approvals. Clinical phase cost averages and standard deviations for the approved drugs in the sample are shown in Table 5. For comparative purposes, the results for the full sample are also shown. Except for phase I, clinical phase costs are notably higher for the approved drugs than for drugs as a whole. Phase II and III costs for the approved drugs are 77 and 18% higher, respectively. This result is qualitatively consistent with what we found in our previous study. An explanation that we offered therein may still be appropriate. The results may reflect a tendency to prioritize development by directing more resources, possibly by conducting more studies concurrently, to investigational drugs that appear, after early testing, to be the most likely to be approved. Since we are not linking

Table 5
Out-of-pocket clinical period phase costs for approved compounds (in millions of 2000 dollars)[a]

Testing phase	Approved drugs[b]			Full sample[c]		
	Mean cost	Median cost	Standard deviation	Mean cost	Median cost	Standard deviation
Phase I	15.2	11.7	14.3	15.2	13.9	12.8
Phase II	41.7	31.5	30.2	23.5	17.0	22.1
Phase III	115.2	78.7	95.0	86.3	62.0	60.6
Long-term animal	4.4	0	5.4	5.2	3.1	4.8

[a] All costs were deflated using the GDP Implicit Price Deflator.
[b] Estimates for the approved drugs are based on data for 24 of the 68 sample drugs.
[c] Weighted values were used in calculating means, medians, and standard deviations for the full sample.

failures to successes here and since we have full phase cost data for the 24 approved drugs, we can add phase costs for each drug to determine a total clinical period cost for each drug and use those data to find confidence intervals for mean out-of-pocket and capitalized clinical period cost for approved drugs. Mean out-of-pocket clinical period cost for the approved drugs was US$ 176.5 million, with a 95% confidence interval of US$ 126–227 million. We used actual phase timing for individual approved drugs, rather than averages over all approved drugs, to capitalize costs for individual approved drugs. Doing so yielded a mean clinical period capitalized cost of US$ 251.3 million, with a 95% confidence interval of US$ 180.2–322.4 million.

The FDA prioritizes new drugs by therapeutic significance at the time of submission of an application for marketing approval.[30] New drugs are rated as either priority (P) or standard (S).[31] Kaitin and Healy (2000), Kaitin and DiMasi (2000), Reichert (2000), and DiMasi (2001a) contain numerous analyses of development and approval times by FDA therapeutic rating. However, the only prior analysis of development costs by therapeutic rating was in our previous study. We found higher mean clinical phase costs for more highly rated drugs. The results for this sample also show higher costs. Mean clinical period out-of-pocket cost for approved drugs with a P rating was US$ 207 million, compared to US$ 155 million for drugs that had received the S rating.

The differential was less for capitalized costs. Mean clinical period capitalized cost was US$ 273 million for drugs with a P rating and US$ 236 million for those with the S rating. In both cases, the confidence intervals for P and S rated drugs overlap. However, given the substantial variability in drug development costs and the fact that the number of compounds in each category was small (10 drugs with a P rating and 14 with an S rating), this outcome is not surprising. However, it is plausible that, on average, testing a priority-rated drug breaks more new scientific ground and so is costlier, as firms must learn through experience. It may also be the case that firms have the incentive to do more wide-ranging and costly testing on drugs that have the potential to be both clinically and commercially significant. Our results can then be viewed as supportive, but not conclusive, evidence of higher costs for drugs with higher therapeutic significance ratings.

7.2. *Cost of post-approval R&D*

Our main objective was to estimate pre-approval R&D costs. However, our pre-approval estimates together with other pharmaceutical industry data regarding the drug development process allowed us to construct an estimate of the cost of post-approval R&D, and thereby obtain an estimate of average total R&D cost per new drug covering the entire

[30] The process is intended to provide direction for internal prioritization of marketing approval reviews by the FDA. The *Prescription Drug User Fee Act of 1992* and its reauthorization in 1997 include performance goals for the FDA that are defined in terms of the therapeutic ratings.

[31] In late 1992 the FDA switched from a three-tiered rating system (A, B, C) to the current two-tiered system (P, S). Drugs that were rated A were judged to represent a significant gain over existing therapy, those rated B were judged to represent a moderate gain over existing therapy, and those rated C were judged to represent little or no gain over existing therapy. Our sample includes drugs that were rated under the old system. We assigned drugs that had received an A or B rating to the P category, and drugs that had received a C rating to the S category under the current system.

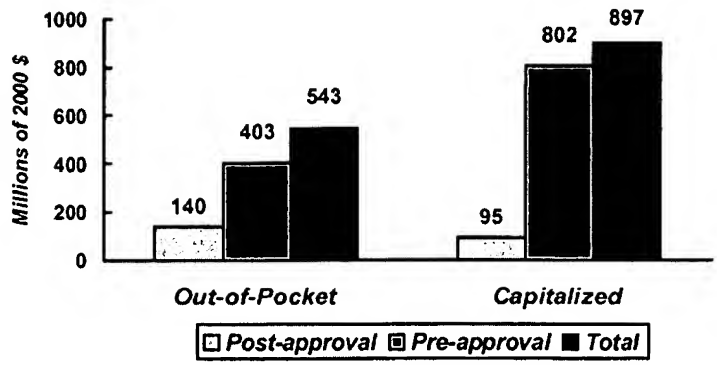


Fig. 4. Out-of-pocket and capitalized total cost per approved new drug for new drugs and for improvements to existing drugs.

development and marketing life-cycle. The aggregate annual expenditure data that we collected for the cost survey firms show that these firms spent approximately three-quarters of their prescription pharmaceutical R&D expenditures on self-originated new drugs, 10% on investigational drugs that are licensed-in or otherwise acquired, and 15% on improvements to drugs that have already been approved.

We cannot, however, use the percentage of aggregate R&D expenditures spent on post-approval R&D on a current basis and apply it to our pre-approval cost estimate to obtain an estimate of the cost of post-approval R&D per approved drug. The reason is that pre-approval costs occur years before post-approval costs. We may use our aggregate annual firm R&D data, but we must build in a reasonable lag structure. Our methodology for doing so is discussed in detail in an appendix that is available from the authors upon request (Appendix A).

We used a 10-year lag for the aggregate data (approximate time between median pre-approval development costs and median post-approval costs), assumed that post-approval R&D cost per approval is the same, on average, for licensed-in and self-originated drugs, and determined the percentage of approvals for the cost survey firms that are self-originated to estimate the ratio of post-approval R&D cost per approved drug to pre-approval cost per approved drug. The data indicated that this share was 34.8%. Thus, we estimated the out-of-pocket cost per approved drug for post-approval R&D to be US$ 140 million (Fig. 4). Since these costs occur after approval and we are capitalizing costs to the point of marketing approval, our discounted cost estimate is lower (US$ 95 million). Thus, out-of-pocket cost per approved drug for post-approval R&D is 25.8% of total R&D cost (pre- and post-approval), while capitalized cost for post-approval R&D is 10.6% of total cost.

7.3. Tax analysis

The cost estimates that are presented here are pre-tax. As noted above, OTA (1993) used the basic data and methodology from our previous study in their report, but the OTA also reported an after-tax figure determined by subtracting a percentage of pre-tax capitalized cost. The percentage was an assumed average effective corporate income tax rate for the

period. Hence, a straightforward calculation can be made to use our R&D cost estimates as inputs in after-tax analyses of R&D rates of return (OTA, 1993; Grabowski and Vernon, 1994). However, some have suggested that an after-income tax figure is the relevant measure of pharmaceutical industry R&D cost (Public Citizen, 2001).

As a stand-alone estimate for R&D cost, we find such a figure to be inadequate for our purposes and potentially misleading. First, we are primarily interested in trends in private sector resource costs associated with getting a new drug to regulatory marketing approval. Tax rates and tax structures can change over time, so trends in resource costs can be masked by after-tax figures. Second, even if the objective is to measure the effective cost to companies, that cost is not properly measured by subtracting the corporate income tax deduction for R&D from the resource cost estimate. It can also be misleading, as it may suggest that government is subsidizing corporate R&D by the amount of the deduction. The corporate income tax is intended to be a tax on profits. Deductions for R&D and other business costs are the means used to approximate the appropriate base for the tax (revenues minus costs). Thus, cost deductions on corporate income tax statements cannot be properly viewed as tax breaks.

The only potential tax advantage with respect to administration of the corporate income tax involves the timing of tax payments. R&D is an investment, but firms are allowed to deduct R&D costs (excluding plant and equipment) as current expenses in lieu of depreciating these investment costs over time. Nevertheless, the value of this timing effect should be significantly less than the total deduction.[32] The accounting informational requirements needed to appropriately depreciate an intangible asset such as R&D are so formidable that expensing of R&D is allowed under accounting guidelines. The true economic depreciation schedule likely varies significantly by industry, by firms within an industry, and by project within a firm. Thus, the practice of allowing what is in effect a 100% depreciation rate in the first year can be viewed as a second-best solution for an otherwise intractable issue.

A portion of the US tax code that is intended to serve as a stimulus to innovation by effectively subsidizing R&D is the Research and Experimentation (R&E) tax credit. The R&E tax credit was not relevant to a significant degree to the study period for our previous analysis (DiMasi et al., 1991). However, it is almost fully applicable to the study period for the current analysis. The credit is generally determined as a percentage of the *excess* of qualified R&D expenditures in a year over a base amount. It is difficult to adequately assess the quantitative impact of this tax policy. Over the history of the implementation of the R&E tax credit, the percentage credited has changed, as has the method for determining

[32] In theory, optimal administration of the tax would involve depreciating all forms of intangible capital at economically appropriate rates. However, tax savings relative to the theoretical optimum should be measured in a tax revenue-neutral context. If intangible capital were depreciated rather than expensed, then the present value of tax revenues would increase. To keep revenues constant, the tax rate would have to be lowered. If all industries were identical with respect to the degrees to which they utilized intangible capital of all types, then tax burdens would not be any different in the alternative state (abstracting from any induced secondary effects on the distribution of industry allocations between tangible and intangible capital or between labor and capital). The pharmaceutical industry, however, is almost certainly above-average in terms of investment in intangible capital (Clarkson, 1977). If the optimal state is attainable at reasonable cost, the tax savings to the pharmaceutical industry, then, is not the difference in the present values of its tax burden as between the current state and the optimum at the current tax rate, but something less that depends on the extent to which the pharmaceutical industry is above-average with regard to investment in intangible capital.

the base amount.[33] It seems unlikely, though, that the credit has had a substantial economic impact on large multinational pharmaceutical firms.[34]

Since 1983 an orphan drug tax credit has also been available to manufacturers for clinical trial expenses related to the development of drugs for orphan indications (fewer than 200,000 patients afflicted in the United States or where it can be demonstrated that development is not profitable). However, for a number of reasons the empirical significance of this credit for the type of firm surveyed for this study is likely to be very small.[35] Analysis of data provided in a Congressional Research Service (CRS) report indicates that orphan drug tax credits amount to a fraction of a percent of pharmaceutical industry R&D expenditures (Guenther, 1999).[36]

[33] In the early implementation years the credit percentage was 25%, but that was lowered to 20% in 1987. The base amount had been an average of research expenditures that met certain criteria for the three previous tax years. In most instances it now essentially involves applying an historical R&D-sales ratio (any 5 years from 1983 to 1988) to the average of gross receipts for the previous 4 tax years. The credit can be applied only to the excess of current "qualified research expenses" over the base amount. A variety of R&D expenditures are excluded from consideration. For example, management expenses other than first-line supervision of those directly engaged in research activity, some computer software development costs, and 35% of research expenses contracted out to for-profit firms are not counted. The credit also does not apply to research conducted outside the United States, Puerto Rico, or any possession of the United States. In addition, firms will typically elect to reduce the allowed credit by the maximum corporate income tax rate (currently 35%). If they do not, then they must reduce the research expenses that they deducted on their corporate income tax statements by the amount of the credit.

[34] Many firms do not separately report R&E tax credits in their published financial data. We did find R&E credits reported in the public financial statements of seven large pharmaceutical firms for each year from 1999 to 2001 (GlaxoSmithKline, Johnson & Johnson, Lilly, Pfizer, Pharmacia, Schering-Plough, and Takeda) and for 2001 for American Home Products (now Wyeth). We compared the credit amounts to the firms' reported R&D expenses. R&E credits as a percentage of R&D expenditures varied somewhat by firm and year (0–5.2%). Overall, the tax credits amounted to 2.0% of R&D expenditures. Adding Merck, which reported on a broader category (General Business Credits), increased the share only to 2.1%. One might argue that prescription pharmaceutical R&D could contribute more to the accumulation of R&E tax credits than is indicated by these data. This might be so if prescription pharmaceutical R&D expenditures grow more rapidly than the firms' other R&D expenditures (this effect would be mitigated, though, in the long-run if pharmaceutical sales also increase at a rate that is greater than for the firms' other businesses). We do not know if this has been the case. However, even if it has, that impact could be more than reversed if firms have made greater use of outsourcing in pharmaceutical than in non-pharmaceutical R&D. By all accounts, pharmaceutical firms have contracted out drug development activities at a rapidly growing rate over our study period, and the share of pharmaceutical R&D expenditures currently accounted for by outsourcing is substantial. As noted above, a significant share of outsourced R&D is excluded from the tax credit calculations.

[35] Unless it can be demonstrated that it is necessary to go outside the United States to find patients, the credit (50% of qualified clinical trial expenses) is not available for foreign trial costs. It is also cannot be applied to clinical testing on any non-orphan indications for a compound with an orphan drug designation. In addition, the vast majority of the manufacturers with products that have received orphan drug designations are biotech firms or small niche pharmaceutical firms (see http://www.fda.gov/orphan/designat/list.htm). For development as a whole, it is highly likely therefore that the share of R&D expenditures for which the orphan drug credit was applicable for traditional large multinational pharmaceutical firms is quite low.

[36] The report includes data on both orphan drug tax credits and taxable income for the pharmaceutical industry for 1990–1994. The CRS also noted in its report that 20.3% of US pharmaceutical industry domestic sales and exports were spent on R&D in 1997. Applying this R&D-sales ratio to the data on taxable income suggests that orphan drug credits amounted to 0.3% of R&D expenditures. This is a conservative estimate for large pharmaceutical firms since taxable income is determined by deducting business expenses from sales, and since, as noted above, biotechnology and small pharmaceutical firms obtain a disproportionate share of the credits.

8. Validation

In their 1993 report, the OTA reviewed the literature on pharmaceutical R&D costs. In addition to critiquing the methodologies used in these studies, the review addressed evidence on the reasonableness of the studies, particularly the DiMasi et al. (1991) study. The OTA concluded that, "the estimates by DiMasi and colleagues of the cash outlays required to bring a new drug to market and the time profile of those costs provide a reasonably accurate picture of the mean R&D cash outlays for NCEs first tested in humans between 1970 and 1982" (OTA, 1993, p. 66). The OTA provided varied data and analyses to corroborate the results in DiMasi et al. (1991). We corroborate the basic cost results in this study by examining the representativeness of our sample firms and by analyzing various independently derived results and data about the industry and the drug development process. We pay particular attention to data that corroborate the growth in costs between the previous study and the current one.

8.1. Internal validation

The Tufts CSDD database of investigational compounds, from which our sample was selected, contains data on the vast majority of new drugs developed in the United States (DiMasi, 2001a). The distribution of investigational drugs across therapeutic classes for our 10 survey firms is very close to the distribution for all drugs in the database. We examined the data for eight specific therapeutic classes and one miscellaneous class for drugs in the database that met study inclusion criteria. There are 530 compounds in the database that meet these criteria and for which a therapeutic class could be identified (272 of these compounds belong to the 10 cost survey firms). The largest difference in share for a specific class between all firms in the database and the cost survey firms was 1.5%.[37] Using a chi-squared goodness-of-fit test comparing the therapeutic class distributions for the cost survey firms and the other firms in the database, we found no statistically significant difference for the class shares ($\chi^2 = 5.01$, d.f. $= 9$).[38]

Based on publicly available data, we also found that pharmaceutical R&D expenditure growth rates for the survey firms as a whole were similar to the reported growth rates for all PhRMA member firms. For example, the annual growth rate in real pharmaceutical R&D expenditures for the survey firms[39] from 1995 to 2000 is 11.3%, compared to 11.0% for PhRMA member firms over the same period.[40]

[37] The population shares for the analgesic/anesthetic, antiinfective, antineoplastic, cardiovascular, central nervous system, endocrine, gastrointestinal, immunologic, miscellaneous, and respiratory classes are 9.1, 12.8, 9.4, 23.2, 17.9, 7.0, 2.1, 3.0, 9.4, and 6.0%, respectively. The corresponding shares for the cost survey firms are 9.6, 14.3, 8.1, 22.8, 19.1, 7.4, 2.2, 3.3, 7.7, and 5.5%, respectively.

[38] The estimated clinical success rate for all firms in this dataset (21.5%) is also very close to the estimated success rate for the 10 firms using the same inclusion criteria (22.2%).

[39] The data are for nine of the 10 firms. We did not find pharmaceutical R&D data for one of the firms, but this firm has a relatively small pharmaceutical subsidiary whose inclusion would not materially affect the results. The data were taken from Scrip's Pharmaceutical Company League Tables (various years) and company annual reports.

[40] The annual growth rate for 1995–1999 was slightly lower for the survey firms compared to all PhRMA member firms (11.5% versus 11.8%).

8.2. External validation

Publicly available data that were collected independently can be examined to determine the extent to which they are consistent with our results in terms of levels or rates of change. Specifically, we examined independent information on clinical trial sizes, measures of clinical trial complexity, and published trade association data on R&D employment and expenditures.

8.2.1. Clinical trial sizes and complexity

Several groups have compiled data on clinical trial sizes for new molecular entities approved in the United States for periods that range from the late 1970s to 2001 (BCG, 1993; OTA, 1993; Peck, 1997; PAREXEL, 2002).[41] Averaging the BCG results for 1981–1984 and 1985–1988 (2277) and comparing them to average of the Peck (1997) and PAREXEL (2002) results for 1994–1995 and 1998–2001 (5603) yields an annual growth rate in clinical trial sizes of 7.47% per year.[42] We may approximate the increases in cost per subject over time by examining the excess of medical care inflation over general price inflation. The medical care component of the CPI increased at an average annual rate of 6.73% from 1984 to 1997, while general price inflation (applying the price index used to deflate costs for this study) rose at an annual rate of 3.06% over the same period. Thus, other things being equal, these results suggest an increase of 11.4% per year in clinical trial costs. This compares to our finding of an 11.8% annual growth rate in out-of-pocket clinical period cost between DiMasi et al. (1991) and the current study.

These separate estimations need not be in perfect agreement because our clinical cost figures include costs not directly related to the number of clinical trial subjects (infrastructure costs, fixed costs related to production of clinical trial supplies, animal testing during the clinical period, etc.). In addition, there could be some economies of scale in clinical testing that would result in a somewhat lower growth in cost per subject. However, data compiled by DataEdge, LLC (PAREXEL, 2002, p. 96) indicate that the complexity of clinical trials

[41] Each of these sources obtained data for a sample of the US approvals during specific periods. The BCG found the mean number of subjects included in NDAs to be 1576 for 1977–1980, 1321 for 1981–1984, and 3233 for 1985–1988. OTA (1993) compared clinical trial sizes for NDAs for three therapeutic categories (antihypertensives, antimicrobials, and nonsteroidal antiinflammatories) over two periods. In aggregate, it found the mean number of subjects to be 2019 for 1978–1983 approvals ($n = 28$) and 3128 for 1986–1990 approvals ($n = 25$). Peck (1997) found the mean number of subjects to be 5507 for 12 of 50 1994–1995 approvals. PAREXEL (2002) has examined the number of subjects in NDAs for 55% of the new molecular entities approved by the FDA in each year from 1998 to 2001. For the period as whole, the mean number of subjects is 5621 ($n = 64$). The latter two averages are similar to what we have found as the mean number of subjects across all three clinical phases for the investigational drugs in our cost survey (5303). CMR (2000) found the mean number of subjects to be 4478 for 23 marketing approval applications submitted from 1995 to 2000. However, only nine of the submissions were to the FDA, with the remainder submitted to European Union and Japanese regulatory authorities. Since pre-approval costs are measured here up to the point of US regulatory approval, we use the US-based data.

[42] These groupings were chosen so that the mean approval years were 1984 and 1997 (the average approval years for the DiMasi et al. (1991) and the current cost samples). The difference in the two periods was taken to be 12.5 years. For the early period, we prefer the BCG data to the OTA data, since the OTA data apply to only three therapeutic categories that likely tend, in aggregate, to have above-average clinical trial sizes.

has increased significantly in recent years. Their index of clinical trial complexity[43] for phases I–III increased at an annual rate of 4.8% per year from 1992 to 2000. An increase in clinical trial complexity will contribute to even higher growth rates for clinical costs.[44]

8.2.2. Growth in industry R&D employment costs

Despite rapid growth in outsourcing of R&D activities over the last few decades, pharmaceutical firms have significantly expanded the number of their own employees devoted to the R&D function. In its industry profile and annual survey reports over various years, PhRMA has provided annual information on the R&D employment of its member firms. From 1980 to 2000, total R&D employment increased at a compound annual rate of 5.4%, with scientific and professional staff increasing at a 7.4% annual rate.[45]

We adjusted National Science Foundation (NSF) data on median annual salaries for full-time employed biological scientists with doctorates working in for-profit life sciences industries from 1993 to 1999 for inflation (GDP Implicit Price Deflator).[46] Real salaries increased at a rate of 1.75% per year over this period. The OTA presented similar data for every 2 years from 1973 to 1989 (OTA, 1993; pp. 62–63). The real growth rate in median annual salaries for biological scientists with doctorates employed in business or industry from 1981 to 1989 was 1.77%. Applying a real growth rate of 1.76% per year for compensation to a growth rate of 7.4% per year in employment yields a growth rate of 9.3% per year for labor costs. This is moderately higher than the growth rate of 7.6% per year that we found for total out-of-pocket cost per approved drug between our previous study and the current one.[47] Thus, some labor costs have grown fairly rapidly. Most of the growth in labor costs, though, has been due to increasing the labor force devoted to R&D, rather than to increases in real wages.

[43] The index is based on the mean number of medical procedures to be applied to patients in clinical trial protocols. Some of these procedures will be covered by insurance, but the index should provide at least a rough indicator of the degree to which the clinical trial process is increasing in complexity.

[44] DataEdge has also compiled information on certain clinical trial costs (investigator fees and central laboratory costs). Changes in cost due to increases in clinical trial complexity will be at least partially reflected in these data. PAREXEL (2002) reports their index of mean costs per subject across all clinical phases (I–IV) for each year from 1996 to 2000. The index increased at an average annual real rate of 5.33% over this period. Combining this growth rate with the above growth rate for clinical trial sizes suggests a 13.1% average annual real rate of increase in clinical trial costs. Piecing together the index values for years reported in earlier editions of PAREXEL (2002) yields a 3.54% real growth rate for 1993–2000. This would imply an 11.2% average annual real growth rate in clinical trial costs.

[45] Over our study period, highly trained personnel have comprised an increasingly large component of the pharmaceutical industry in-house R&D labor force. The share of total R&D personnel for the scientific and professional category in the PhRMA data increased from 56.3% in 1980 to 81.8% in 2000.

[46] The data were compiled for 1993, 1995, 1997, and 1999 by the NSF through surveys of doctoral scientists and engineers in the United States (National Science Foundation, various years). The NSF used a new survey instrument for 1993 and later. Data for every 2 years from 1973 to 1989 used somewhat different occupational definitions. Thus, these data may not be strictly comparable to the data for 1993 and beyond. Data were not available for 1991.

[47] The NSF survey data for 1993–1999 show a real increase of 1.2% per year in median annual salaries across all degrees for biological scientists working in the for-profit life sciences industries. Applying this growth rate to the growth rate of 5.4% for all pharmaceutical industry R&D personnel yields an increase of 6.7% per year in labor costs.

8.2.3. Cost estimates from published industry R&D expenditures

PhRMA has gathered information on aggregate industry R&D spending for decades. The resultant R&D expenditure time series can be linked to data on new drug approvals to develop rough estimates of out-of-pocket pharmaceutical R&D costs. As noted above, linking current expenditures to current approvals is an inadequate approach. Our estimated time profile for a representative drug and the pattern of costs over that timeline determined for this study can be used to construct a lag structure for aggregate expenditures and approvals.[48]

There are two complications regarding the PhRMA data that must be addressed before we can validate our estimates. One is that while PhRMA has traditionally disaggregated its reported R&D expenditure data into expenditures on new drugs and expenditures on improvements to existing drugs, it has not gathered information on how expenditures on new drugs can be further decomposed into expenditures on self-originated and on licensed-in new drugs. Our R&D cost estimates are for self-originated drugs, and a substantial portion of the R&D expenditures on licensed-in drugs are likely missing from the PhRMA data.[49] Thus, we need to associate lagged industry expenditures on self-originated new drugs with self-originated new drug approvals. The second complication is that, with the exception of 1 year, PhRMA has gathered information on the domestic expenditures of all its firms, but the foreign expenditures of only its US-owned members. Our method for dealing with these complications is described in detail in an appendix available from the authors upon request (Appendix B).

We related estimated lagged PhRMA member firm R&D expenditures on self-originated new drugs from 1978 to 1998 to the number of self-originated new drug approvals by PhRMA member firms from 1990 to 2000. The lag structure follows the phase time-expenditure profile implied by our data, with weights attached to aggregate expenditures over a 2–12 year period. The ratio of total lagged self-originated R&D expenditures to the total number of self-originated approvals yields an estimate of the out-of-pocket cost of new drug

[48] PhRMA also publishes a breakdown of annual R&D expenditures of its member firms by function (PhRMA, 2001). The share for the category "Clinical Evaluation: Phases I–III" in 1999 is 29.1%. This share cannot be compared to the clinical period share of total out-of-pocket cost per approved drug implied by our estimates for at least three reasons. First, clinical period costs in a given year are linked to pre-human R&D expenditures in past years, and the pharmaceutical R&D expenditure series shows substantial growth. Thus, shares based on current year expenditures will significantly understate the clinical portion. Second, portions or all of some categories are for expenditures on post-approval R&D and should be deducted from the base before a pre-approval clinical share is computed. For example, given their definitions, the categories for "Clinical Evaluation: Phase IV (11.7%)" and "Process Development for Manufacturing and Quality Control (8.3%)" would likely have to be taken entirely out of the base. In addition, portions of other categories also likely are associated with post-approval R&D. Third, our notion of clinical period costs extends beyond direct patient costs and includes fixed infrastructure costs and other costs incurred during the clinical period. The categories "Toxicology and Safety Testing (4.5%)," Pharmaceutical Dosage Formulation and Stability Testing (7.3%)," "Regulatory: IND and NDA (4.1%)," "Bioavailability (1.8%)," and "Other (9.0%)" would each have to be decomposed into shares for pre-human R&D, pre-approval clinical period R&D, and post-approval R&D. With a reasonable pre-human/clinical lag structure, it is possible to choose an allocation of the three periods for these functional categories that results in a clinical period share of pre-approval R&D expenditures that equals our estimated cost share. However, we are not aware of any data that allows one to make these allocations credibly. Thus, we concluded that the PhRMA data on functional categories could not be used as an external check on our results.

[49] The PhRMA data apply to member firms. Not every pharmaceutical firm (particularly foreign firms) and few biotechnology firms are members of the organization.

development.[50] We calculated a range for this ratio by using reported domestic industry R&D expenditures for a lower bound and domestic plus foreign (inclusive of estimates for foreign-owned firms) industry R&D expenditures as an upper bound. The result is a range of US$ 354–558 million for out-of-pocket cost per approved new drug (inclusive of failures). Our out-of-pocket cost estimate of US$ 403 million per approved drug calculated from our survey data falls within this range. Capitalizing the aggregate expenditure data using our phase-expenditure time profile yields a range of US$ 650–1023 million, which encompasses our total capitalized cost estimate of US$ 802 million.

We also conducted a check similar to what the OTA had done in its report (OTA, 1993, pp. 61–62). In theory, under our average development and approval time profile described above, all industry self-originated new drug R&D expenditures in 1988 would be associated with new drug approvals from 1990 to 2000. If each self-originated new drug approval from 1990 to 2000 by a PhRMA member firm is assumed to cost US$ 403 million, then we can use the yearly time-expenditure weights noted above to estimate PhRMA member firm total self-originated R&D expenditures in 1988. Doing so yields US$ 6176 million in 2000 dollars. This value fits within our range for self-originated new drug R&D expenditures estimated from the PhRMA data (US$ 4942–7777 million in 2000 dollars).

9. Conclusions

The cost of developing new drugs is a topic that has long engendered considerable interest. The interest has intensified recently as firms have become increasingly concerned about improving productivity in a period of consolidation and cost containment pressures in the marketplace, and industry critics question industry statements about the level of R&D costs and the impact that price regulation would have on R&D (Public Citizen, 2001).[51] We have undertaken the only comprehensive project-based analysis of the costs of drug development since our previous study (DiMasi et al., 1991). In the last study we estimated average R&D cost to be US$ 231 million in 1987 dollars. For our updated analysis, we estimated that total R&D cost per new drug is US$ 802 million in 2000 dollars. Our results were validated in an number of ways through analyses of independently derived published data on the pharmaceutical industry. Including an estimate of the cost per approved new drug for R&D conducted after approval increases total R&D cost to nearly US$ 900 million. Our pre-approval estimate represents a two and one-half-fold increase in real capitalized costs. On an annualized basis, the growth rate in inflation-adjusted cost was 7.6% for out-of-pocket expenditures and 7.4% for capitalized costs.

Roughly speaking, the current study covers R&D costs that yielded approvals, for the most part, during the 1990s. The previous study (DiMasi et al., 1991) generally involved

[50] We believe that aggregating over the expenditure and approval periods is superior to using an average of yearly ratios. Year-to-year ratios are highly variable since they are very sensitive to the denominator value (number of self-originated new drug approvals) for the year.

[51] Pre-approval R&D expenditures are sunk costs at the time a pricing decision has to be made. Thus, they should not affect price setting in an unregulated market. However, to the extent that high past R&D costs predict high future R&D costs, then anticipated or realized stringent price regulation can significantly reduce incentives to innovate and thereby negatively impact future drug development.

R&D for 1980s approvals, and the first study in this series (Hansen, 1979) was mainly relevant to 1970s approvals. While the compound annual growth rates in out-of-pocket costs between successive studies were similar (7.0% per year between the first two studies and 7.6% per year between the last two), the rates of increase for the two major R&D phases were quite different. Although both preclinical and clinical period costs increased in real terms in this study, the rate of increase for the preclinical period was less than one-third that for the first two studies, while the growth rate for clinical costs was nearly twice as high for the two most recent studies.

Our data do not allow us to test hypotheses about factors that affect how costs change over time, but some conjectures can be made. For example, over the periods analyzed the pharmaceutical industry has increasingly focused on developing treatments for chronic and degenerative diseases or conditions associated with those diseases.[52] Therapies for such conditions are generally more costly to test, as they typically require more complex patient care and monitoring, longer periods for effects to be observed, or larger trial sizes to establish their efficacy.

When the study periods analyzed for the previous study and the current one are compared, one observes that the number of new drugs approved increased over time, as did the number of drugs investigated. This can be associated with patient recruitment that is more time-consuming and costlier.

Finally, the development of more stringent cost containment strategies in the United States and abroad such as tiered formularies and the demand for cost-effectiveness results may have led firms to test their drugs more often against competitor products already on the market (F-D-C Reports, 1999). This will generally be costlier than testing against placebo; the trials will likely need to be more highly powered (i.e. clinical trial sizes will have to be higher) to establish a statistical difference.

These factors help explain the growth in clinical period costs. Preclinical (discovery and preclinical development) costs also grew in real terms, but much more slowly than in the past. The widespread use of discovery technologies, such as combinatorial chemistry techniques and high-throughput screening, during the current study period may have created enough efficiency gains to slow down the growth of preclinical costs.

The cost growth rates that we have observed are substantial. There is no guarantee that they will continue at these levels, but we can determine where costs would end up if they did. The average approval date for our sample was in 1997. Assuming the same growth rates for out-of-pocket and capitalized costs, then the out-of-pocket pre-approval cost per approved drug for R&D relevant to approvals in 2001 would be US$ 540 million, while capitalized pre-approval cost would be US$ 1.1 billion. If growth rates were maintained and R&D was initiated in 2001 with approvals obtained 12 years later, then pre-approval out-of-pocket cost would rise to US$ 970 million and pre-approval capitalized cost would rise to US$ 1.9 billion.

A number of technical factors can work to alter the growth pattern for future R&D costs. We observed improved clinical phase attrition rates for the current study. If firms

[52] We have in mind a broader concept than chronic use drugs. The conditions treated may require drugs that are used on a short-term, medium-term, or intermittent basis. These conditions may result from the natural course of a chronic disease or they may occur as side effects from direct treatment of such complex diseases.

can further improve their performance in terminating research early for compounds that will not make it to approval, then this will help lower out-of-pocket and capitalized costs. Reductions in development times, other things being equal, would also lower capitalized costs. Some recent evidence on clinical development times suggest a shortened process, at least in the United States (Kaitin and DiMasi, 2000; DiMasi, 2001a), but it is too soon to conclude that we are observing a new trend. Finally, emerging discovery and development technologies may have profound effects on R&D productivity. Industry analysts that have recently examined the impact that genomics and other new technologies may have on the R&D process have suggested that as pharmaceutical firms increasingly embrace the new approaches, R&D costs may actually rise significantly in the short run (Pharma Marketletter, 2001; Tollman et al., 2001). The major reason is that the new technologies may generate many targets that are currently not well understood. Eventually, though, they argue that the science knowledge base will expand sufficiently so that efficiencies will be realized.

Analyses of private sector R&D costs provide a crucial input to policy-oriented studies. For example, R&D cost estimates can be utilized in studies that aim to measure the ex-post profitability of new drug development for a given period. This is a timely issue given recent media attention on R&D productivity issues and problems in the R&D pipelines of many leading firms (Pollack, 2002). Results from our prior studies have in fact been used in analyses of the rate of return to pharmaceutical R&D (Grabowski and Vernon, 1990; OTA, 1993; Grabowski and Vernon, 1994).[53] These studies of the profitability of new drug development have not found evidence of significant and sustained excess profits. The estimated internal rates of return are quite close to the cost-of-capital. The much higher R&D cost estimates for this study raise a question about the recent profit experience of the pharmaceutical industry. However, Grabowski and Vernon (2000) found substantial growth in pharmaceutical sales for 1990s drug cohorts. A new study (Grabowski et al., 2002) on pharmaceutical profitability using some of the cost results in this study and recent sales data is qualitatively consistent with the outcomes of the earlier profitability studies (i.e. the internal rate of return is close to the industry cost-of-capital).

Data on R&D costs can also be helpful in analyzing the impact on R&D returns from policy changes that affect the intellectual property protection system, drug development times, or FDA approval times, and therefore influence private incentives to innovate. The Congressional Budget Office (CBO), for example, examined the net effect on pharmaceutical industry returns that the *Drug Price Competition Act of 1984* had from simultaneously reducing the cost of generic entry and increasing effective patent lifetimes (CBO, 1998). Simulations of proposed policy changes for these and other variables that affect the costs of and returns to pharmaceutical R&D can similarly be conducted using our new estimates.

The relationship between pharmaceutical industry profitability and investment in R&D has recently been examined in Scherer (2001). The author found a high degree of correlation between the deviations from trend for the time series on pharmaceutical industry R&D expenditures and on gross margins, indicating that R&D outlays are affected significantly

[53] As noted above, tax issues are explicitly considered in such studies. The corporate income tax, however, plays a very limited role in such analyses. The reason is that the tax essentially enters symmetrically in the analysis (applied to revenues as well as costs), and so the impact on the internal rate of return is minimal. The net present value of profits, though, is lower because of the tax.

by changes in profitability. The growth rate for gross margins for recent years was also substantially lower than the growth rate for R&D outlays, leading to the suggestion that R&D growth rates could lessen in the future. If that were to happen, one might ask what would happen to R&D costs. This would depend on the outcome of internal rate of return analyses by firms on marginal projects.[54] The ultimate impact on future costs, however, will also depend on whether and to what degree currently unforeseen biomedical advances that expand scientific opportunities will be realized.

Finally, our results indicate that variability in drug development costs has declined somewhat but is still substantial. For an earlier period, DiMasi et al. (1995a) found varying average clinical period costs for a number of major therapeutic classes. We will examine costs by therapeutic category in future research. For that same earlier period, DiMasi et al. (1995b) also found that average R&D costs tended to decrease with firm size. The structure of the traditional pharmaceutical industry appears to have evolved somewhat since then. Examining new drug output levels by firm, DiMasi (2000) found both a long-term deconcentration trend for the research-based pharmaceutical industry and substantial new entry during the 1990s with respect to traditional small molecule output. The R&D cost data for this study can be used in further analyses of R&D productivity at the firm level in future research.

Acknowledgements

We thank two anonymous referees and the editors for helpful comments. All errors and omissions are the responsibility of the authors. We also thank the surveyed firms for providing data, and individuals in those firms who kindly gave their time when we needed some of the responses clarified. The authors did not receive any external funding to conduct this study.

References

Boston Consulting Group, 1993. The Contribution of Pharmaceutical Companies: What's at Stake for America. The Boston Consulting Group, Boston, MA, September 1993.

US Congressional Budget Office, 1998. How Increased Competition from Generic Drugs has Affected Prices and Returns in the Pharmaceutical Industry. US Government Printing Office, Washington, DC, July 1998.

Clarkson, 1977. Intangible Capital and Rates of Return, American Enterprise Institute, Washington, DC.

CMR, 2000. Describing Dossiers: Characterising Clinical Dossiers for Global Registration. R&D Briefing 25, CMR International, Surrey, UK.

DiMasi, J.A., Hansen, R.W., Grabowski, H.G., Lasagna, L., 1991. Cost of innovation in the pharmaceutical industry. Journal of Health Economics 10, 107–142.

DiMasi, J.A., Hansen, R.W., Grabowski, H.G., Lasagna, L., 1995a. Research and development costs for new drugs by therapeutic category: a study of the US pharmaceutical industry. PharmacoEconomics 7, 152–169.

[54] While one might postulate that higher cost projects would be more vulnerable, firms should take account of expected profitability. Given that we found some evidence of higher costs for more innovative products, if firms elect to focus more on innovative projects on expected profitability grounds, average costs would increase when economically marginal projects are dropped.

DiMasi, J.A., Grabowski, H.G., Vernon, J., 1995b. R&D costs, innovative output and firm size in the pharmaceutical industry. International Journal of the Economics of Business 2, 201–219.

DiMasi, J.A., 2000. New drug innovation and pharmaceutical industry structure: trends in the output of pharmaceutical firms. Drug Information Journal 34, 1169–1194.

DiMasi, J.A., 2001a. New drug development in the United States 1963–1999. Clinical Pharmacology & Therapeutics 69, 286–296.

DiMasi, J.A., 2001b. Risks in new drug development: approval success rates for investigational drugs. Clinical Pharmacology & Therapeutics 69, 297–307.

F-D-C Reports, 1999. NDA Submissions are Shrinking in Size but Increasing in Complexity. The Pink Sheet 61, p. 28.

Grabowski, H.G., Vernon, J., 1990. A new look at the returns and risks to pharmaceutical R&D. Management Science 36, 804–821.

Grabowski, H.G., Vernon, J., 1994. Returns to R&D on new drug introductions in the 1980s. Journal of Health Economics 13, 383–406.

Grabowski, H.G., Vernon, J., 2000. The distribution of sales revenues from pharmaceutical innovation. PharmacoEconomics 18 (Supplement 1), 21–32.

Grabowski, H.G., Vernon, J., DiMasi, J.A., 2002. Returns on research and development for 1990s new drug introductions. PharmacoEconomics 20 (Supplement 3), 11–29.

Guenther, G., 1999. Federal taxation of the drug industry from 1990 to 1996. Memorandum to Joint Economic Committee, US Congress, Congressional Research Service, 13 December 1999.

Hansen, R.W., 1979. The pharmaceutical development process: estimates of current development costs and times and the effects of regulatory changes. In: Chien, R.I. (Ed.), Issues in Pharmaceutical Economics. Lexington Books, Lexington, MA, pp. 151–187.

Ibbotson Associates, 2001. Stocks, Bonds, Bills & Inflation: 2001 Yearbook, Ibbotson Associates, Chicago, Illinois.

Kaitin, K.I., Healy, E.M., 2000. The new drug approvals of 1996, 1997, and 1998: drug development trends in the user fee era. Drug Information Journal 34, 1–14.

Kaitin, K.I., DiMasi, J.A., 2000. Measuring the pace of new drug development in the user fee era. Drug Information Journal 34, 673–680.

Myers, S.C., Shyam-Sunder, L., 1996. Measuring pharmaceutical industry risk and the cost-of-capital. In: Helms, R.B. (Ed.), Competitive Strategies in the Pharmaceutical Industry. American Enterprise Institute, Washington, DC, pp. 208–237.

Myers, S.C., Howe, C.D., 1997. A life-cycle financial model of pharmaceutical R&D. Working Paper, Program on the Pharmaceutical Industry. Massachusetts Institute of Technology, Cambridge, MA.

National Institutes of Health, 2000. NIH Response to the Conference Report Request for a Plan to Ensure Taxpayers' Interests are Protected: A Plan to Ensure Taxpayers' Interests are Protected. National Institutes of Health, Rockville, MD, July 2001.

US Congress, Office of Technology Assessment, 1993. Pharmaceutical R&D: Costs, Risks, and Rewards, OTA-H-522. US Government Printing Office, Washington, DC.

Peck, C.C., 1997. Drug development: improving the process. Food and Drug Law Journal 52, 163–167.

Pharma Marketletter, 2001. Genomics may increase costs of NCE development, says Lehman Bros. Pharma Marketletter, vol. 28, 26 February 2001, pp. 24–25.

PhRMA, 2000. Pharmaceutical Industry Profile 2000: Research for the Millennium, Pharmaceutical Research and Manufacturers of America, Washington, DC.

PhRMA, 2001. Pharmaceutical Industry Profile 2001, Pharmaceutical Research and Manufacturers of America, Washington, DC.

PJB Publications Ltd., 2000. Scrip's 2000 Pharmaceutical Company League Tables, Richmond, Surrey, UK, 3 November 2000.

Pollack, A., 2002. Despite Billions for Discoveries, Pipeline of Drugs is Far from Full. New York Times, 19 April 2002.

Public Citizen, 2001. Rx R&D Myths: The Case Against the Drug Industry's R&D Scare Card, Public Citizen, Congress Watch, July 2001.

Reichert, J.M., 2000. New biopharmaceuticals in the US: trends in development and marketing approvals 1995–1999. Trends in Biotechnology 18, 364–369.

Scherer, F.M., 2000. The pharmaceutical industry. In: Newhouse, J.P. (Ed.), Handbook of Health Economics, vol. 1. Elsevier, Amsterdam, Chapter 25, pp. 1297–1336.

Scherer, F.M., 2001. The link between gross profitability and pharmaceutical R&D spending. Health Affairs 20, 216–220.

Tollman, P., Guy, P., Altshuler, J., Flanagan, A., Steiner, M., 2001. A Revolution in R&D: The Impact of Genomics. The Boston Consulting Group, Boston, MA, June 2001.

US Food and Drug Administration, Center for Drug Evaluation and Research, 1999. From Test Tube to Patient: Improving Health Through Human Drugs (special report). US Government Printing Office, Washington, DC, September 1999 (http://www.fda.gov/fdac/special/newdrug/ndd_toc.html).

EXHIBIT M

Huijsduijnen et al., Expert Opin. Ther. Targets, 2002, 6:624

Review

Expert Opinion

Ashley Publications
www.ashley-pub.com

Cardiovascular, Renal, Endocrine and Metabolic

Protein tyrosine phosphatases as drug targets: PTP1B and beyond

Rob Hooft van Huijsduijnen, Sébastien Wälchli, Mark Ibberson & Axel Harrenga
Serono Pharmaceutical Research Institute, 14, chemin des Aulx, Plan-les-Ouates/Geneva, Switzerland

Protein tyrosine phosphatases (PTPs) control signal transduction pathways and have recently emerged as potential drug targets. Inhibition of individual PTPs can result in the activation of therapeutically relevant kinase cascades. This is particularly useful in cases where disease is associated with hormonal resistance, such as insensitivity to insulin or leptin. Currently, PTP1B is being investigated by a number of companies as a promising target for leptin/insulin mimetics and in the treatment of diabetes and obesity. Since all 90 – 100 PTPs have been identified in the human genome, the challenge now is to identify the function of these enzymes and the therapeutic indications that may exist for specific PTP inhibitors.

Keywords: cancer, diabetes, inflammation, obesity, osteoporosis, protein tyrosine phosphatase (PTP)

Expert Opin. Ther. Targets (2002) 6(6):637-647

1. Introduction

Nearly all intracellular signalling is controlled by protein phosphorylation/dephosphorylation events. The role of kinases as phosphorylating enzymes was appreciated very early on, while the active role of their counterparts, the phosphatases, is now also rapidly gaining recognition. The dynamic character of intracellular phosphorylation and dephosphorylation reactions is best appreciated when cells are treated with generic phosphatase inhibitors, such as vanadate or arsenate, which results in the rapid and massive phosphorylation of many intracellular targets and which has pleiotropic physiological effects [1]. One may anticipate from this observation that inhibition of individual phosphatases would stimulate individual, specific pathways; several examples of this concept are provided below. The authors will focus on protein tyrosine phosphatases (PTPs), the largest group of phosphatases, as potential drug targets and summarise current knowledge of PTP1B, which is the paradigm PTP target for insulin- and leptin-mimetic drugs.

2. Protein tyrosine phosphatase gene families

Tyrosine phosphatases can be subdivided into 'classical PTPs', which are phosphotyrosine specific, and 'dual-specific PTPs' (DS-PTPs), which also dephosphorylate serine- and threonine-phosphate residues. Classical and DS-PTPs are characterised by the PTP signature motif ([H/V]C(X)$_5$R[S/T]). Weaker sequence similarity extends to an ~ 250 amino acid domain [2,3]. These PTPs form the bulk of the tyrosine phosphatases but the signature motif alone is also found in other, smaller subfamilies whose overall sequence similarity with these two main classes is not obvious. These include Cdc25A-C, LMW-PTPs (low molecular weight-PTPs), the MTM (myotubularin)-related and the tensin (TPTPE)-like phosphatases. The authors have explored the Celera and NCBI (National Center for Biotechnology Information) draft human genome sequences for tyrosine phosphatase-like genes based on sequence similarities to each of these subfamilies, based on the presence of

Table 1. The number of PTP family genes predicted from Celera and NCBI-generated human genome sequences, based on the various PTP subclasses.

	Celera		NCBI	
Profile	*Genes*	*Pseudogenes*	*Genes*	*Pseudogenes*
'Classical' PTPs	39	13	35	12
DS-PTPs	33	5	31	4
Cdc25a-c	3	0	3	0
LMW-PTPs	1	3	1	2
MTM-like	14	4	14	4
Tensin (TPTPE)-like	5	7	5	7
Total	95	32	89	29

DS-PTP: Dual-specific PTP; LMW-PTP: Low molecular weight PTP; MTM: Myotubularin; NCBI: National Center for Biotechnology Information; PTP: Protein tyrosine phosphatase; TPTPE: Tensin-like PTP.

the signature motif and intron–exon predictions (Table 1). For a number of these phosphatases and even for several classical and DS-PTPs, it is uncertain as to whether they possess innate or inducible enzymatic activity [4]. Other phosphatases such as PTEN (phosphatase and tensin homologue deleted on chromosome ten) and the MTM family members have phospholipids as substrates [5]. Interestingly, the number of classical and DS-PTPs predicted from the genome is not dramatically different from an earlier survey based on cDNA and EST (expressed sequence tag) data [2]. A very similar observation was made for another drug target class [501]. From a drug target discovery point of view, the availability of the human genome sequence does not appear to provide many new short-term opportunities. The accumulation of PTP family member sequences has failed to keep pace with their functional understanding and the authors believe that the PTPs are underexploited as drug targets.

Consistent with a specific role for PTPs in vertebrate growth factor and cytokine signalling, almost no PTPs are found in bacteria (except in pathogenic species, as discussed below). A recent survey of the *Arabidopsis thaliana* plant genome reported only a single PTP out of 112 phosphatase genes [6].

3. Protein tyrosine phosphatases in signalling

Figure 1 represents a simplified view of signalling by receptor kinases (green) and PTPs (red, with catalytic domains represented as rectangles). Typically, cytokine and growth factor receptor signalling is initiated by ligand-induced dimerisation, which activates the receptor's kinase domain. Alternatively, receptors may recruit intracellular kinases such as Jaks (Janus-activated kinases) that phosphorylate the receptor in addition to themselves. Both events trigger phosphorylation cascades. It is increasingly clear that multiple negative feedback mechanisms exist that modulate these pathways including: receptor internalisation and degradation; the induction of SOCS (suppressors of cytokine signalling), which bind and mask receptor phosphotyrosines; PIAS (protein inhibitors of activated STATs [signal transducers and activators of transcription]); and dephosphorylation of cascade components by PTPs (see [7] for a review of cytokine signalling suppressors). Since PTPs can exist as either soluble protein or membrane-bound 'receptor' PTPs associated with the endoplasmic reticulum, dephosphorylation can take place while the receptor is membrane-anchored or while it is being recycled [8].

These feedback mechanisms are likely to be complex and subtle. For instance, SOCS are transcriptionally activated by cytokines; the PTP, SHP-1 (Src-homology-2 domain-containing phosphatase-1), is post-translationally activated by phosphorylation and can then dock via its Src-homology-2 (SH-2) domain on receptor phosphotyrosines. PTPs and their potential receptor substrates have recently been reviewed [9,10]. A number of PTPs have also been shown to regulate second messenger phosphoproteins. A number of these intracellular targets are listed in Table 2.

4. Protein tyrosine phosphatase drug targets

In this section the authors review PTPs that may be considered as promising drug targets. These include PTPs that are already widely recognised as such by the industry: PTP1B, CD45 and Cdc25B. However, the authors also include a number of more speculative targets; the choice for this second set has an inherent element of subjectivity.

4.1 PTP1B: a paradigm for protein tyrosine phosphatases as 'druggable' targets

Since the end of the nineteenth century, vanadate and its derivatives have been known to have therapeutic utility in diabetes [1]. By 1989, it was clear that this effect was mediated by the metal's inhibition of a phosphatase, which might critically regulate the insulin receptor's phosphorylation state [51,52]. Although PTP1B was one of the first PTPs to be cloned, its use as a drug target was only appreciated when it

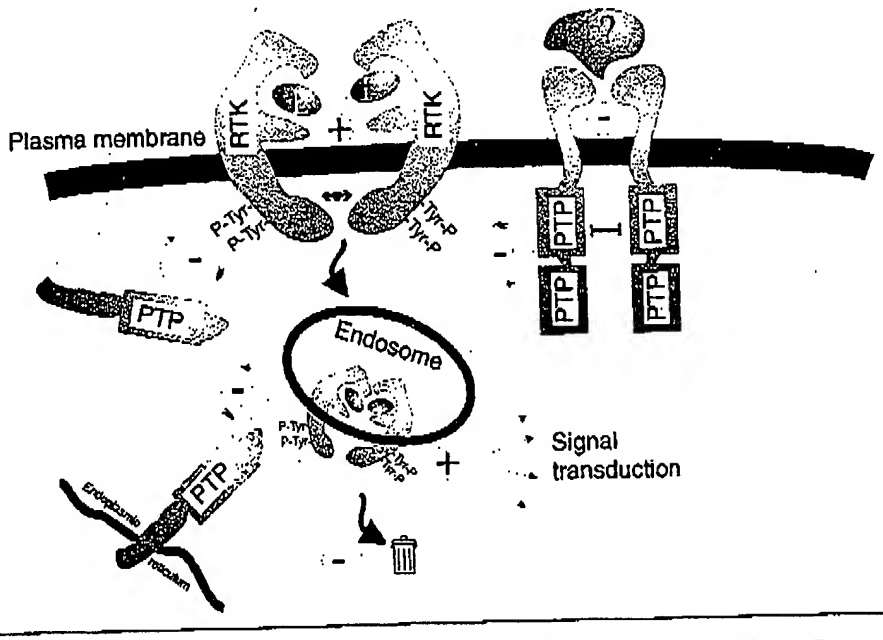


Figure 1. Schematic view of receptor tyrosine kinases and PTPs in signalling. Green: Receptor kinase; Orange: Receptor- and non-receptor-PTPs; PTP catalytic domains are rectangles; Blue: Receptor kinase ligand; Pink: (Putative) receptor PTP ligand. L: Ligand; PTP: Protein tyrosine phosphatase; RTK: Receptor tyrosine kinase.

was discovered that PTP1B knockout mice display reduced plasma glucose and insulin levels, obesity-resistance and increased insulin receptor and IRS-1 (insulin receptor substrate-1) tyrosine phosphorylation [53,54]. These findings have incited many companies to start PTP1B drug screening projects. The development of an early candidate (Ertiprotafib®, Wyeth) was discontinued following a diabetes Phase II trial because of efficacy and side effect issues. However, the prospects for PTP1B-targeting treatments are still bright, especially following impressive results from once-weekly injections of antisense reagents by Isis Pharmaceuticals [119], currently licensed to Merck; PTP1B is considered 'an outstanding target' by experts for these disease areas (B Kahn, as cited in [56]).

While PTP1B is widely recognised as a drug target, its precise mode of action is not entirely clear. The enzyme has good substrate specificity for the autophosphorylated insulin receptor [57], but recently an additional, inhibitory role in leptinR signalling was reported, presumably via Jak-2 dephosphorylation [58]. These recent findings help explain certain aspects of the knockout mice, such as their increased energy expenditure and obesity resistance.However, attributing the relative contributions of these two pathways to the PTP1B knockout phenotype is complicated by the fact that leptin and insulin have overlapping functions. Insulin receptors in the brain are involved in the control of obesity and appetite, and enhanced signalling through the leptin receptor reduces adiposity, which results in improved glucose tolerance.

4.2 Protein tyrosine phosphatases in the immune system

4.2.1 Protein tyrosine phosphatases that control antigen receptor signalling

The role of PTPs in the immune system has recently been reviewed [90-95]. One of the earliest PTPs to be discovered, CD45, is essential for T and B lymphocyte antigen receptor signalling [59,60]. This PTP would therefore appear to be a good target for immunosuppressors. The SH-2-domain-containing PTP, SHP-1 (not to be confused with SHIP [SH-2-domain-containing inositol 5-phosphatase], a phospholipid phosphatase), is also involved in T-cell receptor (TCR) and B cell signalling, but its knockout phenotype [83-85] suggests it plays a much wider role, which perhaps makes it less suitable as a drug target.

4.2.2 Protein tyrosine phosphatases in infectious diseases

There are many interesting examples in which pathogens exploit PTPs to increase their survival. The earliest known example is the *Yersinia* bacterium, which encodes a PTP (YopH) that is essential for virulence *in vivo* [96,97]. *Helicobacter pylori*, a common, infectious bacterium and a causative agent of gastric ulcers and stomach cancer, is known to transduce the CagA protein into the gastric epidermal cells on which it thrives. Recently, it was found that one of CagA's roles is to activate SHP-2 (Src-homology-2 domain-containing phosphatase-2) [98]. Another example is *Salmonella*, which is known to transduce a PTP (SptP) into its host cells [99]. Other bacteria (*Mycobacteria, Salmonella*) are also suspected to

Table 2. PTPs associated with intracellular phosphoprotein cascades.

Substrate	Associated PTP	Method used	Ref.
p130^{cas}	PTP-PEST	Trapping mutants and overexpression	[11,12]
	PTP-α	Overexpression of PTP	[13]
	LAR	Overexpression of wt and mutant constructs	[14,15]
	PTP1B	Isolation from immunoprecipitation	[16,17]
	SHP-2	Overexpression	[18]
	PTP-SAP		[19]
STAT-1	SHP-1	Knockout and overexpression	[20,21]
	SHP-2	Direct or indirect mutagenesis	[22]
	PTP-ε	Overexpression and mutagenesis	[23]
STAT-3	SHP-2	Direct or indirect mutagenesis	[22]
	SHP-1	Overexpression	[21]
	PTP-ε	Overexpression and mutagenesis	[24]
STAT-5a/b	SHP-2	Substrate trapping assay, phosphatase assay	[25,26]
	PTP1B	Overexpression and substrate trapping	[27]
	TC-PTP	Overexpression and substrate trapping	[28]
STAT-6	SHP-1	Transgenic expression	[29]
p70zap	CD-45	Knockout in cells and *in vitro* assay	[30]
	HPC-PTP	Transgenic expression	[31]
p52Shc	TC-PTP	Trapping mutants	[32]
c-Abl	PTP-PEST	Knockout cells	[33]
Jak-2	CD-45	Knockout in cells and *in vitro* assay	[34]
	PTP1B	Overexpression, knockout and trapping mutant	[35]
IRS-1/2	PTP1B	Trapping mutants and *in vitro* assay	[36]
	LAR	Transgenic overexpression	[37]
Src-family kinases	PTP-α	Overexpression	[38-40]
	CD-45	Overexpression and cellular knockout	[41,42]
	cPTP-λ, STEP	*In vitro* and *in vivo* assays	[43,44]
	PTP-D1	Overexpression	[45]
	SHP-2	Overexpression and *in vitro* assay	[46]
	HPC-PTP	Cellular assay	[31]
	PTP1B	Cellular knockout and *in vitro* assay	[47,48]
	PTP-BL	*In vitro* assay, negative regulator	[49]
β-catenin	PTP-ζ (mRPTP-β)	*In vitro* and ligand-induced inactivation	[50]

For receptor substrates see [9].
HPC-PTP: Homologue of PC12 cell PTP; IRS: Insulin receptor substrate; Jak: Janus-activated kinase; LAR: Leucocyte antigen-related PTP; PTP: Protein tyrosine phosphatase; PTP-PEST: PEST domain-containing PTP; PTP-SAP: Stomach cancer-associated-PTP; SHP: Src-homology-2 domain-containing phosphatase; Src: Sarcoma rous virus oncogene; STAT: Signal transducer and activator of transcription; STEP: Striatum-enriched phosphatase; TC-PTP: T cell-PTP; wt: Wildtype.

manipulate their hosts with PTPs [98]. It is by no means clear if these intracellular or PTP-transducing microorganisms target the same host pathway(s) but such PTPs would *a priori* seem to be good drug targets.

More indirect evidence of pathogens using PTPs comes from the observation that an established drug for the treatment of leishmaniasis, sodium stibogluconate, is in fact a PTP inhibitor with selectivity towards SHP-1 and, to a lesser extent, SHP-2 [100]. Thus, SHP-1, SHP-2 and these microbial PTPs appear to be effective targets in the treatment of infectious diseases.

4.3 Vascular leakage

Vascular endothelial monolayers play an important role in the control of inflammation. Local inflammation involves cytokine-induced upregulation of endothelial adhesion molecules such as L- and E-selectins and enhanced permeability of tight junctions, followed by neutrophil and macrophage extravasation. It was recently shown that angiopoietin-1 and its endothelial receptor, Tie-2, antagonise this process, i.e., they restore blood vessel integrity from leakiness and play an anti-inflammatory role [101]. It was also shown that endothelial-specific PTP-β (or VE-PTP [vascular endothelial-PTP] for the murine orthologue) specifically dephosphorylates activated Tie-2 receptor kinase *in vitro* and *in vivo* [102]. This would suggest that PTP-β is a drug target in inflammation, with inhibitors blocking neutrophil and macrophage extravasation.

Finally, genetic ablation of the PTP-ε gene results in macrophage abnormalities, notably in response to lipopolysaccharide [103]. This would suggest that PTP-ε is a potential target in septic shock.

Table 3. PTP knockout mice and human PTP mutation phenotypes.

PTP	Knockout/mutated phenotype	Ref.
CD45	Lack T cells, immature B cells*	[59,60]
Cdc25b	Meiotic arrest*[1]	[61]
Cdc25c	No obvious phenotype*	[62]
DEP/CD148/PTPβ2	Die at mid-gestation with severe defects in vascular organisation[3]; frequently deleted in human cancers‡	[63]
EPM2A2/Lafora	Progressive myoclonus epilepsy‡	[64,65]
GLEPP-1	Reduced renal filtration surface area*	[66]
LAR	Mammary gland defect; reduced plasma glucose and insulin levels*	[67] [68]
LC-PTP (He-PTP)	No obvious phenotype*	[69]
MKP-1	No obvious phenotype*	[70]
MTM-1	Human mutations: severe hypotonia and generalised muscle weakness*[2]	[71]
MTMR-2	Human mutations: Charcot–Marie–Tooth Type 4B disease (CMT4B), progressive de-myelination*	[72]
PEZ/PTP36	Androgenisation of female mice, no mammary gland tissue*	Patent WO 02/45500 A2
PTEN	Early embryonic lethality; frequently mutated in human cancers‡	[73,74]
PTP1B	Enhanced leptin and insulin sensitivity*	[53,54]
PTP-PEST	Die at embryonal stage day 8.5, defective neural fold closure, lack primary hepatocytes‡	ML Tremblay, pers. commun.
PTP-α	No obvious phenotype*	[75]
PTP-δ	Growth retardation, early mortality, posture and motor defects‡	[76]
PTP-ε	Hypomyelination*	[77]
PTP-IA-2	Abnormal glucose-stimulated insulin secretion*	[78]
PTP-κ	No obvious phenotype*	[79]
PTP-μ	No obvious phenotype*	Cited in [80]
PTP-σ	Pituitary displasia, defects in olfactory lobes, reduction in CNS size and cell number*	[81]
PTP-ζ	'Suggestion of a fragility of myelin'*	[82]
SHP-1	Moth-eaten (me) mice lethal. Haemopoietic dysregulation, splenomegaly, runting, autoimmune disease‡	[83-85]
SHP-2	Lethal at embryonal stage day 8.5 – 10, defect in mesoderm patterning‡	[86-88]
TC-PTP	Born normal, die 3 – 5 weeks postpartum; defects in both haematopoiesis and immune function‡	[89]

*Healthy, overtly normal, breed normally.
‡Lethal.
[1]Females sterile.
[2]Human mutation.
[3]Manuscript submitted, D Thomas, pers. comm.
GLEPP: Glomerular epithelial protein; LAR: Leucocyte antigen-related PTP; MKP: MAP kinase phosphatase; MTM: Myotubularin; PTEN: Phosphatase and tensin homologue deleted on chromosome ten; PTP: Protein tyrosine phosphatase; PTP-PEST: PEST domain-containing PTP; SHP: Src-homology-2 domain-containing phosphatase; TC-PTP: T cell-PTP.

4.4 Protein tyrosine phosphatase targets in cancer

4.4.1 *Few protein tyrosine phosphatases are tumour suppressors*

It is remarkable, given the large number of kinases whose over-production or increased activity is associated with cancer, that few PTP tumour suppressors are known. An exception is DEP/PTPR-J (density-enhanced PTP), whose mutations were recently found to be associated with colon cancer [63]. Another recent study looked at expression of 68 phosphatase genes in ovarian cancer and found a significant correlation between repression of MKP-1 (MAP kinase phosphatase-1, a DS-PTP) and malignancy [104]. However, no MKP-1 mutations have

(yet) been associated with disease. PTEN, a major tumour suppressor that indirectly inactivates phosphatidylinositol-3 (PI3) kinase [105] is a phospholipid phosphatase rather than a PTP. To the authors' knowledge, none of the 26 or so PTP knockout mice that have been described so far show a tendency to develop neoplasia (Table 3). It is possible that Mother Nature has hedged her bets by providing for redundancy in the dephosphorylation control of critical growth factor-stimulated kinase cascades. By contrast, some PTPs that regulate Src-kinases or Cdc's may be drug targets in cancer.

4.4.2 Protein tyrosine phosphatases that activate Src-family kinases

The activity of Src (sarcoma rous virus oncogene) and other Src-family kinases is controlled by phosphorylation of a C-terminal tyrosine residue. De-phosphorylation of these kinases at this critical residue by a PTP disrupts an intramolecular interaction with Src's N-terminal SH-2 domain. A large number of PTPs have been postulated to dephosphorylate and activate proto-oncogene Src-family kinases (Table 2). It is very doubtful that inhibition of each of these PTPs would block Src overactivity. However, mice lacking PTP-α have reduced Src and fyn activity [75]. Similarly, PTP-ε knockout mice have reduced steady-state levels of activated Src, which suggests that PTP-ε plays a role in dephosphorylating and activating c-Src. When these knockout animals are mated with Neu oncogene-overexpressing animals they produce transgenic offspring with fewer mammary gland tumours than Neu mice that are not mutated in PTP-ε, which suggests that in these animals Src family oncogenes cannot readily cooperate with the Neu oncogene to produce malignancies [106]. A complication is that in many naturally-occurring cancers, the C-terminal autoinhibitory tyrosine of Src is simply deleted or mutated, which frees the oncogene product from PTP control.

4.4.3 Cyclin-dependent kinase-activating protein tyrosine phosphatases

Cdc25s are a family of three PTPs that activate cyclin-dependent kinases [107]. Among these, Cdc25B is overexpressed in a high proportion (> 30%) of primary breast tumours. Cdc25B is being explored by a number of companies as a drug target and inhibitors have shown efficacy in tumour models [108].

4.4.4 A protein tyrosine phosphatase that is specifically required for metastasis formation?

It was recently discovered that Prl-3 is specifically amplified and overexpressed in metastases but not in their originating primary tumours [109]. This elegant work is strongly suggestive of Prl-3 being essential for the formation of metastases but perhaps not for normal tissue growth. However, it remains to be established as to whether Prl-3 activity is merely required for the establishment of metastases or for their survival also. Only in the latter case would a drug be highly effective.

Table 4. PTPs for which structural information is available [55].

PTP	Number of structures	Species
PTP1B	29	Human
TC-PTP	1	Human
SHP-1	2	Human
SHP-2	1	Human
HPC-PTP	1	Mouse
PTEN	1	Human
PTP-α	1	Mouse
PTP-μ	1	Human
LAR	1	Human
Yop51	1	*Yersinia*
SptP	2	*Salmonella*
VHR	2	Human
Kap	2	Human
MKP3/Pyst1	1	Human
Cdc25A	1	Human
Cdc25B	4	Human
LMW-PTP	8	Human, bovine, *Saccharomyces cerevisiae*

LAR: Leucocyte antigen-related PTP; LMW-PTP: Low molecular weight-PTP; PTEN: Phosphatase and tensin homologue deleted on chromosome 10; PTP: Protein tyrosine phosphatase; SHP: Src-homology-2 domain-containing phosphatase; TC-PTP: T cell PTP; VHR: Vaccinia H1-related protein.

4.5 Osteoporosis

Experiments involving antisense PTP-oχ have implicated GLEPP (glomerular epithelial protein)-1 (its human orthologue) as a negative regulator of osteoclastic resorption [110]. These results suggest that blocking GLEPP-1 may have a useful anabolic effect in osteoporosis. Another study observed that alendronate (ALN), an aminobisphosphonate used in the treatment of osteoporosis, is in fact an inhibitor of PTP-ε, a PTP that is highly expressed in osteoclasts [111]. The authors of this study speculate therefore that PTP-ε may be a useful target for the treatment of osteoporosis.

5. Conclusion and expert opinion

5.1 Protein tyrosine phosphatase target discovery strategies

The PTPs that have been reviewed here differ in how well they are established as targets. Some PTPs, like PTP1B, are well accepted as targets, and PTP1B is now the focus of inhibitor discovery at approximately 20 companies [56], whereas for others their 'druggability' is more speculative. Validating a PTP drug target is not a trivial task. Much more

so than for other signalling components, predicting the effects of PTP inhibition from overexpression or substrate specificity studies alone has not been very successful. The history of PTP1B's discovery as a target clearly illustrates that purely deterministic approaches alone are not appropriate for PTPs. Today, the exact substrate of PTP1B is not known (The insulin receptor? IRS-1? Jak-2? The leptin receptor?), but there is very good evidence that PTP1B is a good drug target. It would not be wise to count on the next PTP target to manifest its 'targetability' as clearly as PTP1B. An effective, systematic PTP target approach would instead focus on knockout animals and antisense or other inhibitors as tested in disease models. Expression patterns and the molecular biology of PTPs may provide hints as to disease situations in which these reagents and animals should be tested. The recent revolution in siRNA-based technologies that can be applied in mammalian systems may be very helpful in this respect [112-117].

5.2 A dilemma in new protein tyrosine phosphatase drug target discovery

The authors' current understanding of cell biology in general and of PTPs in particular is, by objective standards, rudimentary. Therefore, they believe that at present a 'knock them out and see what happens' approach is more effective than 'dissective' or 'deterministic' procedures, no matter how elegant or awe-inspiring such latter experiments may be. Unfortunately, 'direct' target validation strategies lack academic and scientific appeal. In addition, the drug industry is not sufficiently investing in efficient, systematic target discovery, because the acquisition of intellectual property for a drug target presently provides an inadequate return of investment. The financial rewards instead come from marketable inhibitors that are subsequently developed against that target (for instance by an industrial competitor). This situation may well be partly responsible for the paucity of new drugs against new targets that enter the market, and it is unlikely that further company mergers are the solution.

5.3 Are protein tyrosine phosphatases tractable targets?

Discovering specific PTP inhibitors is a challenge. For instance, few of the PTP1B inhibitors reported so far in the literature and in patents show impressive selectivity over TC-PTP (T cell-PTP), its nearest homologue, which has very similar substrate selectivity to PTP1B [57]. This is worrisome because TC-PTP knockout mice die soon after birth [89]. On the other hand, hererozygous TC-PTP knockout animals are fertile and normal in all respects, which would suggest that some degree of TC-PTP inhibition may be harmless. Also, vanadate-based compounds, which inhibit virtually all PTPs (and hit quite a few non-PTP targets as well), have been tested in small-scale human trials with only very modest side effects. This suggests that a PTP1B inhibitor need not have absolute selectivity. In the longer term, other PTPs are likely to benefit from the massive screening efforts on PTP1B, through the discovery of new classes of inhibitors [118] and the generation of detailed structural protein information (Table 4). The first PTP cDNAs were cloned in the early nineties, some 3 decades after kinases had been discovered and understood to be implicated in cancers. Yet, the first useful kinase inhibitor drug imatinib mesylate (Gleevec™, Novartis) only entered the market in 2001. A comparison with the kinase timeline suggests that it may be simply too early to tell if and when PTP targeting drugs will be ready for the clinic.

Acknowledgements

Due to space limitations not all original work could be cited. The authors refer to the reviews for further reading and apologise to authors whose work could not be cited for the omission.

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. MORINVILLE A, MAYSINGER D, SHAVER A: From Vanadis to Atropos: vanadium compounds as pharmacological tools in cell death signalling. *Trends Pharmacol. Sci.* (1998) 19(11):452-460.
   - • Pharmacological properties of vanadium-based compounds.
2. HOOFT VAN HUIJSDUIJNEN R: Protein tyrosine phosphatases: counting the trees in the forest. *Gene* (1998) 225(1-2):1-8.
3. ANDERSEN JN, MORTENSEN OH, PETERS GH *et al.*: Structural and evolutionary relationships among protein tyrosine phosphatase domains. *Mol. Cell. Biol.* (2001) 21(21):7117-7136.
   - •• Extensive overview of mammalian PTPs, excellent web support.
4. WISHART MJ, DIXON JE: Gathering STYX: phosphatase-like form predicts functions for unique protein-interaction domains. *Trends BioChem. Sci.* (1998) 23(8):301-306.
   - •• Description and hypotheses for PTPs that are not catalytically active.
5. MAEHAMA T, TAYLOR GS, DIXON JE: PTEN and myotubularin: novel phosphoinositide phosphatases. *Ann. Rev. Biochem.* (2001) 70:247-279.
6. KERK D, BULGRIEN J, SMITH DW, BARSAM B, VERETNIK S, GRIBSKOV M: The complement of protein phosphatase catalytic subunits encoded in the genome of *Arabidopsis*. *Plant Physiol.* (2002) 129(2):908-925.
7. HILTON DJ: Negative regulators of cytokine signal transduction. *Cell. Mol. Life Sci.* (1999) 55(12):1568-1577.
   - • Overview of negative signalling mechanisms.
8. HAJ FG, VERVEER PJ, SQUIRE A, NEEL BG, BASTIAENS PI: Imaging sites of receptor dephosphorylation by PTP1B on the surface of the endoplasmic reticulum. *Science* (2002) 295(5560):1708-1711.
9. ESPANEL X, WÄLCHLI S.

PESCINI GOBERT R *et al.*: Pulling strings below the surface: hormone receptor signaling through inhibition of protein tyrosine phosphatases. *Endocrine* (2001) 15(1):19-28.

10. OSTMAN A, BÖHMER FD: Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatases. *Trends Cell Biol.* (2001) 11(6):258-266.

11. GARTON AJ, FLINT AJ, TONKS NK: Identification of p130(cas) as a substrate for the cytosolic protein tyrosine phosphatase PTP-PEST. *Mol. Cell. Biol.* (1996) 16(11):6408-6418.
- Use of substrate-trapping PTP mutants to identify substrates.

12. GARTON AJ, BURNHAM MR, BOUTON AH, TONKS NK: Association of PTP-PEST with the SH3 domain of p130cas; a novel mechanism of protein tyrosine phosphatase substrate recognition. *Oncogene* (1997) 15(8):877-885.

13. BUIST A, BLANCHETOT C, TERTOOLEN LG, DEN HERTOG J: Identification of p130cas as an *in vivo* substrate of receptor protein-tyrosine phosphatase alpha. *J. Biol. Chem.* (2000) 275(27):20754-20761.

14. WENG LP, WANG X, YU Q: Transmembrane tyrosine phosphatase LAR induces apoptosis by dephosphorylating and destabilizing p130Cas. *Genes Cells* (1999) 4(3):185-196.

15. WANG X, WENG LP, YU Q: Specific inhibition of FGF-induced MAPK activation by the receptor-like protein tyrosine phosphatase LAR. *Oncogene* (2000) 19(19):2346-2353.

16. LIU F, SELLS MA, CHERNOFF J: Transformation suppression by protein tyrosine phosphatase 1B requires a functional SH3 ligand. *Mol. Cell. Biol.* (1998) 18(1):250-259.

17. LIU F, SELLS MA, CHERNOFF J: Protein tyrosine phosphatase 1B negatively regulates integrin signaling. *Curr. Biol.* (1998) 8(3):173-176.

18. MANIE SN, ASTIER A, HAGHAYEGHI N *et al.*: Regulation of integrin-mediated p130(Cas) tyrosine phosphorylation in human B cells. A role for p59(Fyn) and SHP2. *J. Biol. Chem.* (1997) 272(25):15636-15641.

19. NOGUCHI T, TSUDA M, TAKEDA H *et al.*: Inhibition of cell growth and spreading by stomach cancer-associated protein-tyrosine phosphatase-1 (SAP-1) through dephosphorylation of p130cas. *J. Biol. Chem.* (2001) 276(18):15216-15224.

20. DAVID M, CHEN HE, GOELZ S, LARNER AC, NEEL BG: Differential regulation of the alpha/beta interferon-stimulated Jak/Stat pathway by the SH2 domain-containing tyrosine phosphatase SHPTP1. *Mol. Cell. Biol.* (1995) 15(12):7050-7058.

21. TENEV T, BOHMER SA, KAUFMANN R *et al.*: Perinuclear localization of the protein-tyrosine phosphatase SHP-1 and inhibition of epidermal growth factor-stimulated STAT1/3 activation in A431 cells. *Eur. J. Cell. Biol.* (2000) 79(4):261-271.

22. SYMES A, STAHL N, REEVES SA *et al.*: The protein tyrosine phosphatase SHP-2 negatively regulates ciliary neurotrophic factor induction of gene expression. *Curr. Biol.* (1997) 7(9):697-700.

23. TANUMA N, SHIMA H, NAKAMURA K, KIKUCHI K: Protein tyrosine phosphatase epsilonC selectively inhibits interleukin-6- and interleukin-10-induced JAK-STAT signaling. *Blood* (2001) 98(10):3030-3034.

24. TANUMA N, NAKAMURA K, SHIMA H, KIKUCHI K: Protein-tyrosine phosphatase PTPepsilon C inhibits Jak-STAT signaling and differentiation induced by interleukin-6 and leukemia inhibitory factor in M1 leukemia cells. *J. Biol. Chem.* (2000) 275(36):28216-28221.

25. BERCHTOLD S, VOLAREVIC S, MORIGGL R, MERCEP M, GRONER B: Dominant negative variants of the SHP-2 tyrosine phosphatase inhibit prolactin activation of Jak2 (janus kinase 2) and induction of Stat5 (signal transducer and activator of transcription 5)-dependent transcription. *Mol. Endocrinol.* (1998) 12(4):556-567.

26. YU CL, JIN YJ, BURAKOFF SJ: Cytosolic tyrosine dephosphorylation of STAT5. Potential role of SHP-2 in STAT5 regulation. *J. Biol. Chem.* (2000) 275(1):599-604.

27. AOKI N, MATSUDA T: A cytosolic protein-tyrosine phosphatase PTP1B specifically dephosphorylates and deactivates prolactin-activated STAT5a and STAT5b. *J. Biol. Chem.* (2000) 275(50):39718-39726.
- STAT-5 as a target of PTP1B.

28. AOKI N, MATSUDA T: A nuclear protein tyrosine phosphatase TC-PTP is a potential negative regulator of the PRL-mediated signaling pathway: dephosphorylation and deactivation of signal transducer and activator of transcription 5a and 5b by TC-PTP in nucleus. *Mol. Endocrinol.* (2002) 16(1):58-69.

29. HAQUE SJ, HARBOR P, TABRIZI M, YI T, WILLIAMS BR: Protein-tyrosine phosphatase Shp-1 is a negative regulator of IL-4- and IL-13-dependent signal transduction. *J. Biol. Chem.* (1998) 273(51):33893-33896.

30. MUSTELIN T, WILLIAMS S, TAILOR P *et al.*: Regulation of the p70zap tyrosine protein kinase in T cells by the CD45 phosphotyrosine phosphatase. *Eur. J. Immunol.* (1995) 25(4):942-946.

31. RAAB M, RUDD CE: Hematopoietic cell phosphatase (HCP) regulates p56LCK phosphorylation and ZAP-70 binding to T cell receptor zeta chain. *Biochem. Biophys. Res. Comm.* (1996) 222(1):50-57.

32. TIGANIS T, BENNETT AM, RAVICHANDRAN KS, TONKS NK: Epidermal growth factor receptor and the adaptor protein p52Shc are specific substrates of T-cell protein tyrosine phosphatase. *Mol. Cell. Biol.* (1998) 18(3):1622-1634.

33. CONG F, SPENCER S, COTE JF *et al.*: Cytoskeletal protein PSTPIP1 directs the PEST-type protein tyrosine phosphatase to the c-Abl kinase to mediate Abl dephosphorylation. *Mol. Cell* (2000) 6(6):1413-1423.

34. IRIE-SASAKI J, SASAKI T, MATSUMOTO W *et al.*: CD45 is a JAK phosphatase and negatively regulates cytokine receptor signalling. *Nature* (2001) 409(6818):349-354.

35. MYERS MP, ANDERSEN JN, CHENG A *et al.*: TYK2 and JAK2 are substrates of protein-tyrosine phosphatase 1B. *J. Biol. Chem.* (2001) 276(51):47771-47774.
- Jak-2 as a PTP1B substrate, rationale for role in leptinR signalling.

36. GOLDSTEIN BJ, BITTNER-KOWALCZYK A, WHITE MF, HARBECK M: Tyrosine dephosphorylation and deactivation of insulin receptor substrate-1 by protein-tyrosine phosphatase 1B. Possible facilitation by the formation of a ternary complex with the Grb2 adaptor protein. *J. Biol. Chem.* (2000) 275(6):4283-4289.

37. ZABOLOTNY JM, KIM YB, PERONI OD *et al.*: Overexpression of the LAR (leukocyte antigen-related) protein-tyrosine phosphatase in muscle causes insulin resistance. *Proc. Natl. Acad. Sci. USA* (2001) 98(9):5187-5192.

38. ZHENG XM, WANG Y, PALLEN CJ: Cell transformation and activation of pp60c-src by overexpression of a protein tyrosine phosphatase. *Nature* (1992) 359(6393):336-339.

39. DEN HERTOG J, PALS CE, PEPPELENBOSCH MP, TERTOOLEN LG, DE LAAT SW, KRUIJER W: Receptor protein tyrosine phosphatase alpha activates pp60c-src and is involved in neuronal differentiation. *EMBO J.* (1993) 12(10):3789-3798.

40. BHANDARI V, LIM KL, PALLEN CJ: Physical and functional interactions between receptor-like protein-tyrosine phosphatase alpha and p59fyn. *J. Biol. Chem.* (1998) 273(15):8691-8698.

41. HURLEY TR, HYMAN R, SEFTON BM: Differential effects of expression of the CD45 tyrosine protein phosphatase on the tyrosine phosphorylation of the lck, fyn, and c-src tyrosine protein kinases. *Mol. Cell. Biol.* (1993) 13(3):1651-1656.

42. CAHIR MCFARLAND ED, HURLEY TR, PINGEL JT, SEFTON BM, SHAW A, THOMAS ML: Correlation between Src family member regulation by the protein-tyrosine-phosphatase CD45 and transmembrane signaling through the T-cell receptor. *Proc. Natl. Acad. Sci. USA* (1995) 90(4):1402-1406.

43. FANG KS, SABE H, SAITO H, HANAFUSA H: Comparative study of three protein-tyrosine phosphatases. Chicken protein-tyrosine phosphatase lambda dephosphorylates c-Src tyrosine 527. *J. Biol. Chem.* (1994) 269(31):20194-20200.

44. NGUYEN TH, LIU J, LOMBROSO PJ: Striatal enriched phosphatase 61 dephosphorylates fyn at phosphotyrosine 420. *J. Biol. Chem.* (2002) 277(27):24274-24279.

45. MOLLER NP, MOLLER KB, LAMMERS R, KHARITONENKOV A, SURES I, ULLRICH A: Src kinase associates with a member of a distinct subfamily of protein-tyrosine phosphatases containing an ezrin-like domain. *Proc. Natl. Acad. Sci. USA* (1994) 91(16):7477-7481.

46. PENG ZY, CARTWRIGHT CA: Regulation of the Src tyrosine kinase and Syp tyrosine phosphatase by their cellular association. *Oncogene* (1995) 11(10):1955-1962.

47. BJORGE JD, PANG A, FUJITA DJ: Identification of protein-tyrosine phosphatase 1B as the major tyrosine phosphatase activity capable of dephosphorylating and activating c-Src in several human breast cancer cell lines. *J. Biol. Chem.* (2000) 275(52):41439-41446.

* Dephosphorylation of Src by PTP1B.

48. CHENG A, BAI. GS, KENNEDY BP, TREMBLAY ML: Attenuation of adhesion-dependent signaling and cell spreading in transformed fibroblasts lacking protein tyrosine phosphatase-1B. *J. Biol. Chem.* (2001) 276(28):25848-25855.

49. PALMER A, ZIMMER M, ERDMANN KS *et al.*: EphrinB phosphorylation and reverse signaling: regulation by Src kinases and PTP-BL phosphatase. *Mol. Cell* (2002) 9(4):725-737.

50. MENG K, RODRIGUEZ-PENA A, DIMITROV T *et al.*: Pleiotrophin signals increased tyrosine phosphorylation of beta beta-catenin through inactivation of the intrinsic catalytic activity of the receptor-type protein tyrosine phosphatase beta/zeta. *Proc. Natl. Acad. Sci. USA* (2000) 97(6):2603-2608.

** Demonstration that pleiotrophin is a ligand of PTP-ζ.

51. FANTUS IG, KADOTA S, DERAGON G, FOSTER B, POSNER BI: Pervanadate [peroxide(s) of vanadate] mimics insulin action in rat adipocytes via activation of the insulin receptor tyrosine kinase. *Biochemistry* (1989) 28(22):8864-8871.

52. JACKSON TK, SALHANICK AI, SPARKS JD, SPARKS CE, BOLOGNINO M, AMATRUDA JM: Insulin-mimetic effects of vanadate in primary cultures of rat hepatocytes. *Diabetes* (1988) 37(9):1234-1240.

53. KLAMAN LD, BOSS O, PERONI OD *et al.*: Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice. *Mol. Cell. Biol.* (2000) 20(15):5479-5489.

** Important confirmation of PTP1B knockout phenotype and analysis of leptin-like effects.

54. ELCHEBLY M, PAYETTE P, MICHALISZYN E *et al.*: Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* (1999) 283(5407):1544-1548.

** Landmark PTP1B knockout paper.

55. MØLLER N, IVERSEN L, ANDERSEN HS, MCCORMACK JG: Protein tyrosine phosphatases (PTPs) as drug targets: inhibitors of PTP1B for the treatment of diabetes. *Curr. Opin. Drug Disc. Dev.* (2000) 3(5):527-540.

** Rationale of PTP1B as a drug target, and new inhibitors.

56. BROWER V: Fighting fat: new drugs against obesity in the pipeline. *EMBO Rep.* (2002) 3(7):601-602.

57. WÄLCHLI S, CURCHOD M-L, PESCINI GOBERT R, ARKINSTALL S, HOOFT VAN HUIJSDUIJNEN R: Identification of tyrosine phosphatases that dephosphorylate the insulin receptor: a brute-force approach based on 'substrate-trapping' mutants. *J. Biol. Chem.* (2000) 275(13):9792-9796.

58. COOK WS, UNGER RH: Protein tyrosine phosphatase 1B: a potential leptin resistance factor of obesity. *Dev. Cell* (2002) 2(4):385-387.

* Analysis of new data showing that PTP1B plays a role in leptinR signalling.

59. BYTH KF, CONROY LA, HOWLETT S *et al.*: CD45-null transgenic mice reveal a positive regulatory role for CD45 in early thymocyte development, in the selection of CD4+CD8+ thymocytes, and B cell maturation. *J. Exp. Med.* (1996) 183(4):1707-1718.

60. KISHIHARA K, PENNINGER J, WALLACE VA *et al.*: Normal B lymphocyte development but impaired T cell maturation in CD45-exon6 protein tyrosine phosphatase-deficient mice. *Cell* (1993) 74(1):143-156.

** Paper showing that CD45 is essential for antigen receptor signalling.

61. LINCOLN AJ, WICKRAMASINGHE D, STEIN P *et al.*: Cdc25b phosphatase is required for resumption of meiosis during oocyte maturation. *Nat. Genet.* (2002) 30(4):446-449.

62. CHEN MS, HUROV J, WHITE LS, WOODFORD-THOMAS T, PIWNICA-WORMS H: Absence of apparent phenotype in mice lacking Cdc25C protein phosphatase. *Mol. Cell. Biol.* (2001) 21(12):3853-3861.

63. RUIVENKAMP CA, VAN WEZEL T, ZANON C *et al.*: Ptprj is a candidate for the mouse colon-cancer susceptibility locus Scc1 and is frequently deleted in human cancers. *Nat. Genet.* (2002) 31(3):295-300.

* Identification of DEP as a tumour suppressor.

64. MINASSIAN BA, LEE JR, HERBRICK JA *et al.*: Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. *Nat. Genet.* (1998) 20(2):171-174.

65. GANESH S, DELGADO-ESCUETA AV, SAKAMOTO T *et al.*: Targeted disruption of the Epm2a gene causes formation of Lafora inclusion bodies, neurodegeneration, ataxia, myoclonus epilepsy and impaired behavioral response in mice. *Hum. Mol. Genet.* (2002) 11(11):1251-1262.

66. WHARRAM BL, GOYAL M, GILLESPIE PJ *et al.*: Altered podocyte structure in GLEPP1 (Ptpro)-deficient mice associated with hypertension and low glomerular filtration rate. *J. Clin. Invest.* (2000) 106(10):1281-1290.

67. SCHAAPVELD RQ, SCHEPENS JT, ROBINSON GW *et al.*: Impaired mammary gland development and function in mice lacking LAR receptor-like tyrosine phosphatase activity. *Dev. Biol.* (1997) 188(1):134-146.

68. REN JM, LI PM, ZHANG WR *et al.*: Transgenic mice deficient in the LAR protein-tyrosine phosphatase exhibit profound defects in glucose homeostasis. *Diabetes* (1998) 47(3):493-497.

69. GRONDA M, ARAB S, IAFRATE B, SUZUKI H, ZANKE BW: Hematopoietic protein tyrosine phosphatase suppresses extracellular stimulus-regulated kinase activation. *Mol. Cell. Biol.* (2001) 21(20):6851-6858.

70. DORFMAN K, CARRASCO D, GRUDA M, RYAN C, LIRA SA, BRAVO R: Disruption of the erp/mkp-1 gene does not affect mouse development: normal MAP kinase activity in ERP/MKP-1-deficient fibroblasts. *Oncogene* (1996) 13(5):925-931.

71. LAPORTE J, HU LJ, KRETZ C *et al.*: A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved in yeast. *Nat. Genet.* (1996) 13(2):175-182.

72. BOLINO A, MUGLIA M, CONFORTI FL *et al.*: Charcot-Marie-Tooth Type 4B is caused by mutations in the gene encoding myotubularin-related protein-2. *Nat. Genet.* (2000) 25(1):17-19.

73. DI CRISTOFANO A, PESCE B, CORDON-CARDO C, PANDOLFI PP: PTEN is essential for embryonic development and tumour suppression. *Nat. Genet.* (1998) 19(4):348-355.

74. DAHIA PL: PTEN, a unique tumor suppressor gene. *Endocr. Relat. Cancer* (2000) 7(2):115-129.

75. PONNIAH S, WANG DZ, LIM KL, PALLEN CJ: Targeted disruption of the tyrosine phosphatase PTPalpha leads to constitutive downregulation of the kinases Src and Fyn. *Curr. Biol.* (1999) 9(10):535-538.

- Evidence for a role of PTP-α in activating Src-like proto-oncogenes.

76. UETANI N, KATO K, OGURA H *et al.*: Impaired learning with enhanced hippocampal long-term potentiation in PTPδ-deficient mice. *EMBO J.* (2000) 19(12):2775-2785.

77. PERETZ A, GIL-HENN H, SOBKO A, SHINDER V, ATTALI B, ELSON A: Hypomyelination and increased activity of voltage-gated K(+) channels in mice lacking protein tyrosine phosphatase epsilon. *EMBO J.* (2000) 19(15):4036-4045.

78. SAEKI K, ZHU M, KUBOSAKI A, XIE J, LAN MS, NOTKINS AL: Targeted disruption of the protein tyrosine phosphatase-like molecule IA-2 results in alterations in glucose tolerance tests and insulin secretion. *Diabetes* (2002) 51(6):1842-1850.

79. SKARNES WC, MOSS JE, HURTLEY SM, BEDDINGTON RS: Capturing genes encoding membrane and secreted proteins important for mouse development. *Proc. Natl. Acad. Sci. USA* (1995) 92(14):6592-6596.

80. DEN HERTOG J, BLANCHETOT C, BUIST A, OVERVOORDE J, VAN DER SAR A, TERTOOLEN LG: Receptor protein-tyrosine phosphatase signalling in development. *Int. J. Dev. Biol.* (1999) 43(7):723-733.

81. ELCHEBLY M, WAGNER J, KENNEDY TE *et al.*: Neuroendocrine dysplasia in mice lacking protein tyrosine phosphatase sigma. *Nat. Genet.* (1999) 21(3):330-333.

82. HARROCH S, PALMERI M, ROSENBLUTH J *et al.*: No obvious abnormality in mice deficient in receptor protein tyrosine phosphatase β. *Mol. Cell. Biol.* (2000) 20(20):7706-7715.

83. TSUI HW, SIMINOVITCH KA, DE SOUZA L, TSUI FW: Motheaten and viable motheaten mice have mutations in the haematopoietic cell phosphatase gene. *Nat. Genet.* (1993) 4(2):124-129.

- Discovery of the SHP-1 knockout phenotype.

84. KOZLOWSKI M, MLINARIC-RASCAN I, FENG GS, SHEN R, PAWSON T, SIMINOVITCH KA: Expression and catalytic activity of the tyrosine phosphatase PTP1C is severely impaired in motheaten and viable motheaten mice. *J. Exp. Med.* (1993) 178(6):2157-2163.

85. UMEDA S, BEAMER WG, TAKAGI K *et al.*: Deficiency of SHP-1 protein-tyrosine phosphatase activity results in heightened osteoclast function and decreased bone density. *Am. J. Pathol.* (1999) 155(1):223-233.

86. SAXTON TM, CIRUNA BG, HOLMYARD D *et al.*: The SH2 tyrosine phosphatase SHP2 is required for mammalian limb development. *Nat. Genet.* (2000) 24:420-423.

87. YU DH, QU CK, HENEGARIU O, LU X, FENG GS: Protein-tyrosine phosphatase Shp-2 regulates cell spreading, migration, and focal adhesion. *J. Biol. Chem.* (1998) 273(33):21125-21131.

88. ARRANDALE JM, GORE-WILLSE A, ROCKS S *et al.*: Insulin signaling in mice expressing reduced levels of Syp. *J. Biol. Chem.* (1996) 271(35):21353-21358.

89. YOU-TEN KE, MUISE ES, ITIE A *et al.*: Impaired bone marrow microenvironment and immune function in T cell protein tyrosine phosphatase-deficient mice. *J. Exp. Med.* (1997) 186(5):683-693.

90. NEEL BG: Introduction: phosphatases in the immune system. *Semin. Immunol.* (2000) 12(4):347-348.

- Series of extensive reviews on the role of various PTPs in the immune system.

91. ALEXANDER DR: The CD45 tyrosine phosphatase: a positive and negative regulator of immune cell function. *Semin. Immunol.* (2000) 12(4):349-359.

92. ZHANG J, SOMANI AK, SIMINOVITCH KA: Roles of the SHP-1 tyrosine phosphatase in the negative regulation of cell signalling. *Semin. Immunol.* (2000) 12(4):361-378.

93. IBARRA-SANCHEZ MJ, SIMONCIC PD, NESTEL FR, DUPLAY P, LAPP WS, TREMBLAY ML: The T-cell protein tyrosine phosphatase. *Semin. Immunol.* (2000) 12(4):379-386.

94. SAXENA M, MUSTELIN T: Extracellular signals and scores of phosphatases: all roads lead to MAP kinase. *Semin. Immunol.* (2000) 12(4):387-396.

95. KRYSTAL G: Lipid phosphatases in the immune system. *Semin. Immunol.* (2000) 12(4):397-403.

96. BLISKA JB, GUAN KL, DIXON JE, FALKOW S: Tyrosine phosphate hydrolysis of host proteins by an essential *Yersinia* virulence determinant. *Proc. Natl. Acad. Sci. USA* (1991) 88(4):1187-1191.

97. DEVINNEY I, STEELE-MORTIMER I, FINLAY BB: Phosphatases and kinases delivered to the host cell by bacterial pathogens. *Trends Microbiol.* (2000) 8(1):29-33.

98. HIGASHI H, TSUTSUMI R, MUTO S *et al.*: SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. *Science* (2002) 295(5555):683-686.
- Important due to the role of SHP-1 in infectious diseases.

99. MURLI S, WATSON RO, GALAN JE: Role of tyrosine kinases and the tyrosine phosphatase SptP in the interaction of Salmonella with host cells. *Cell. Microbiol.* (2001) 3(12):795-810.

100. PATHAK MK, YI T: Sodium stibogluconate is a potent inhibitor of protein tyrosine phosphatases and augments cytokine responses in hemopoietic cell lines. *J. Immunol.* (2001) 167(6):3391-3397.
- Indirect evidence for the role of PTPs in infectious disease.

101. JAIN RK, MUNN LL: Leaky vessels? Call Ang1! *Nat. Med.* (2000) 6(2):131-132.

102. FACHINGER G, DEUTSCH U, RISAU W: Functional interaction of vascular endothelial-protein-tyrosine phosphatase with the angiopoietin receptor Tie-2. *Oncogene* (1999) 18(43):5948-5953.
- Evidence for a role of PTP-β in Tie-2 signalling.

103. THOMPSON LJ, JIANG J, MADAMANCHI N, RUNGE MS, PATTERSON C: PTP-ε, a tyrosine phosphatase expressed in endothelium, negatively regulates endothelial cell proliferation. *Am. J. Physiol. Heart Circ. Physiol.* (2001) 281(1):H396-H403.

104. MANZANO RG, MONTUENGA LM, DAYTON M *et al.*: CL100 expression is down-regulated in advanced epithelial ovarian cancer and its re-expression decreases its malignant potential. *Oncogene* (2002) 21(28):4435-4447.

105. MYERS MP, PASS I, BATTY IH *et al.*: The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. *Proc. Natl. Acad. Sci. USA* (1998) 95(23):13513-13518.
- Paper showing that PTEN is a major tumour suppressor.

106. DOMBRÁDI V, KRIEGLSTEIN J, KLUMPP S: Regulating the regulators. Conference on protein phosphorylation and protein phosphatases. *EMBO Rep.* (2002) 3(2):120-124.

107. GALAKTIONOV K, LEE AK, ECKSTEIN J *et al.*: CDC25 phosphatases as potential human oncogenes. *Science* (1995) 269(5230):1575-1577.

108. LAZO JS, NEMOTO K, PESTELL KE *et al.*: Identification of a potent and selective pharmacophore for Cdc25 dual specificity phosphatase inhibitors. *Mol. Pharmacol.* (2002) 61(4):720-728.

109. SAHA S, BARDELLI A, BUCKHAULTS P *et al.*: A phosphatase associated with metastasis of colorectal cancer. *Science* (2001) 294(5545):1343-1346.
- Elegant study suggesting that Prl-3 is required for metastases.

110. SUHR SM, PAMULA S, BAYLINK DJ, LAU KH: Antisense oligodeoxynucleotide evidence that a unique osteoclastic protein-tyrosine phosphatase is essential for osteoclastic resorption. *J. Bone Miner. Res.* (2001) 16(10):1795-1803.

111. SCHMIDT A, RUTLEDGE SJ, ENDO N *et al.*: Protein-tyrosine phosphatase activity regulates osteoclast formation and function: inhibition by alendronate. *Proc. Natl. Acad. Sci. USA* (1996) 93(7):3068-3073.

112. TUSCHL T: Expanding small RNA interference. *Nat. Biotechnol.* (2002) 20(5):446-448.
- Review of new siRNA expression techniques from the laboratory that first made RNAi work reproducibly in mammalian cells.

113. MIYAGISHI M, TAIRA K: U6 promoter driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotechnol.* (2002) 20(5):497-500.

114. LEE NS, DOHJIMA T, BAUER G *et al.*: Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotechnol.* (2002) 20(5):500-505.

115. PAUL CP, GOOD PD, WINER I, ENGELKE DR: Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* (2002) 20(5):505-508.

116. SUI G, SOOHOO C, AFFAR EL B, GAY F, SHI Y, FORRESTER WC: A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. USA* (2002) 99(8):5515-5520.

117. PADDISON PJ, CAUDY AA, BERNSTEIN E, HANNON GJ, CONKLIN DS: Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* (2002) 16(8):948-958.

118. BLASKOVICH MA, KIM HO: Recent discovery and development of protein tyrosine phosphatase inhibitors. *Expert Opin. Ther. Patents* (2002) 12(6):871-905.
- Extensive review of PTP inhibitors from patents and the literature.

119. ZINKER BA, RONDINONE CM, TREVILLYAN JM *et al.*: PTP1B antisense oligonucleotide lowers PTP1B protein, normalizes blood glucose, and improves insulin sensitivity in diabetic mice. *Proc. Natl. Acad. Sci. USA* (2002) 99(17):11357-11362.

Patent

201. DELTAGEN INC.: WO02/45500A2 (2001).

Website

301. http://genomebiology.com/2001/2002/2008/research/0029.2001 MAGLICH JM, SLUDER A, GUAN X *et al.*: Comparison of complete nuclear receptor sets from the human, *Caenorhabditis elegans* and *Drosophila* genomes. *Genome Biol.* (2001) 2(8):RESEARCH0029.

Affiliation

Rob Hooft van Huijsduijnen PhD[1†], Sébastien Wälchli PhD[1], Mark Ibberson PhD[1] & Axel Harrenga PhD[2]
†Author for correspondence
[1]Serono Pharmaceutical Research Institute, 14, chemin des Aulx, Plan-les-Ouates/Geneva, Switzerland
Tel: +41 22 70 696 02; Fax: +41 22 79 469 65; E-mail: rob.hooft@serono.com
[2]Astex Technology Ltd, 250 Cambridge Science Park, CB4 0WE Cambridge, UK

EXHIBIT N

Hopkins & Groom, Nature Reviews Drug Discovery 1:727 (2002)

OPINION

The druggable genome

Andrew L. Hopkins and Colin R. Groom

An assessment of the number of molecular targets that represent an opportunity for therapeutic intervention is crucial to the development of post-genomic research strategies within the pharmaceutical industry. Now that we know the size of the human genome, it is interesting to consider just how many molecular targets this opportunity represents. We start from the position that we understand the properties that are required for a good drug, and therefore must be able to understand what makes a good drug target.

Biological systems contain only four types of macromolecule with which we can interfere using small-molecule therapeutic agents: proteins, polysaccharides, lipids and nucleic acids. Toxicity, specificity and the inability to obtain potent compounds against the latter three types means that the vast majority of successful drugs achieve their activity by binding to, and modifying the activity of, a protein. This limits the molecular targets for which commercially viable compounds can be developed, leading to the concept of 'the druggable genome' — the subset of the ~30,000 genes in the human genome that express proteins able to bind drug-like molecules.

One way of assessing the opportunities available to the pharmaceutical industry is to begin by studying the properties that are required in a commercially viable drug. For the most part, this means an orally bioavailable compound. The physico-chemical properties that are necessary to increase the likelihood of oral bioavailability have been formalized into the 'rule-of-five'[1] (BOX 1). Constraints such as these dictate the type of protein we see as drug targets — simply put, drug targets need to be able to bind compounds with appropriate properties.

Druggable protein families

The druggable subset of the human genome can be predicted using several methods. In a comprehensive review of the accumulated portfolio of the pharmaceutical industry, Drews[2,3] identified 483 targets, and concluded that there could be 5,000–10,000 potential targets on the basis of an estimate of the number of disease-related genes[4]. However, this analysis did not focus on the properties of the drugs that define those targets. The idea of assessing the number of ligand-binding domains has also recently been introduced as a measure of the number of potential points at which small-molecule therapeutic agents could act — suggestions are that this figure could be even greater than 10,000 (REF. 5).

Binding sites on proteins usually exist out of functional necessity; therefore, most successful drugs achieve their activity by competing for a binding site on a protein with an endogenous small molecule. For a drug to be effective, it must bind to its molecular target with a reasonable degree of potency. Our analysis of the Investigational Drugs Database (produced by Current Drugs) and the Pharmaprojects Database (produced by PJB Publications), in addition to a thorough review of the literature, identifies 399 non-redundant molecular targets that have been shown to bind rule-of-five-compliant compounds with binding affinities below 10 μM.

Although there is some degree of overlap with earlier work[2–4], we have captured several proteins that are targeted by experimental drugs, and eliminated some targets for which activity has not yet been shown to be modulated by rule-of-five-compliant compounds. Most of the drugs and leads that were identified in this survey are competitive with an endogenous ligand at a structurally defined binding site.

We have taken the sequences of the drug-binding domains of these proteins and determined the families that they represent, as captured by their InterPro domain[6,7]. Only 130 protein families represent the known drug targets (ONLINE TABLE 1). Nearly half of the targets fall into just six gene families: G-protein-coupled receptors (GPCRs), serine/threonine and tyrosine protein kinases, zinc metallopeptidases, serine proteases, nuclear hormone receptors and phosphodiesterases (FIG. 1a).

Box 1 | **Guidelines for oral bioavailability: the 'rule-of-five'**

The 'rule-of-five' analysis by Lipinski *et al.*[1] shows that poor absorption or permeation of a compound are more likely when: there are more than five hydrogen-bond donors; the molecular mass is more than 500 Da; the lipophilicity is high (expressed as cLogP > 5); and the sum of nitrogen and oxygen atoms is more than 10. These rules, more appropriately described as guidelines, do not cover drugs that are derived from natural products, for which other absorption mechanisms are involved.

Clearly, published data on the oral bioavailability of existing drugs could be used as a method for defining the properties of viable drugs; however, our approach using the rule-of-five allows predictions to be made. In practice, the number of targets identified by applying the rule-of-five filters differs little from that obtained solely by literature analysis of all known drugs, whether rule-of-five compliant or not.

Table 1 | **Comparison of the druggable genomes of selected eukaryotes**

	Homo sapiens	*Drosophila melanogaster*	*Caenorhabditis elegans*	*Saccharomyces cerevisiae*
Total number of predicted genes[8,9,16]	~30,000	13,601	18,424	6,241
Number of proteins in proteome*	21,688	13,849	17,946	6,127
Number of estimated druggable targets	3,051	1,714	2,267	508
Percentage that are predicted druggable targets	~10–14%	12%	12%	8%

Three hundred and seventy-six targets identified to bind rule-of-five-compliant drugs have had InterPro domains assigned. The prevalence of these InterPro domains in various genomes has then been determined. Twenty-three more bacterial and viral drug targets for which InterPro assignments could not be made have not been included in any of our analyses. *Data taken from InterPro, 29 October 2001.

The sequence and functional similarities within a gene family are usually indicative of a general conservation of binding-site architecture between family members. This would suggest that if one member of a gene family were able to bind a drug, other members would also be able to bind a compound with similar physico-chemical properties. Using this reasoning, 3,051 of the predicted 30,000 or so genes in the human genome[8,9] code for a protein with some precedent for binding a drug-like molecule (FIG. 1b, ONLINE TABLE 1). A comparative analysis of the eukaryotic genomes of worm, fly and yeast also reveals that approximately one in ten of the proteins expressed by these genomes belong to gene families with members that have previously shown modulation by small-molecule drugs (TABLE 1).

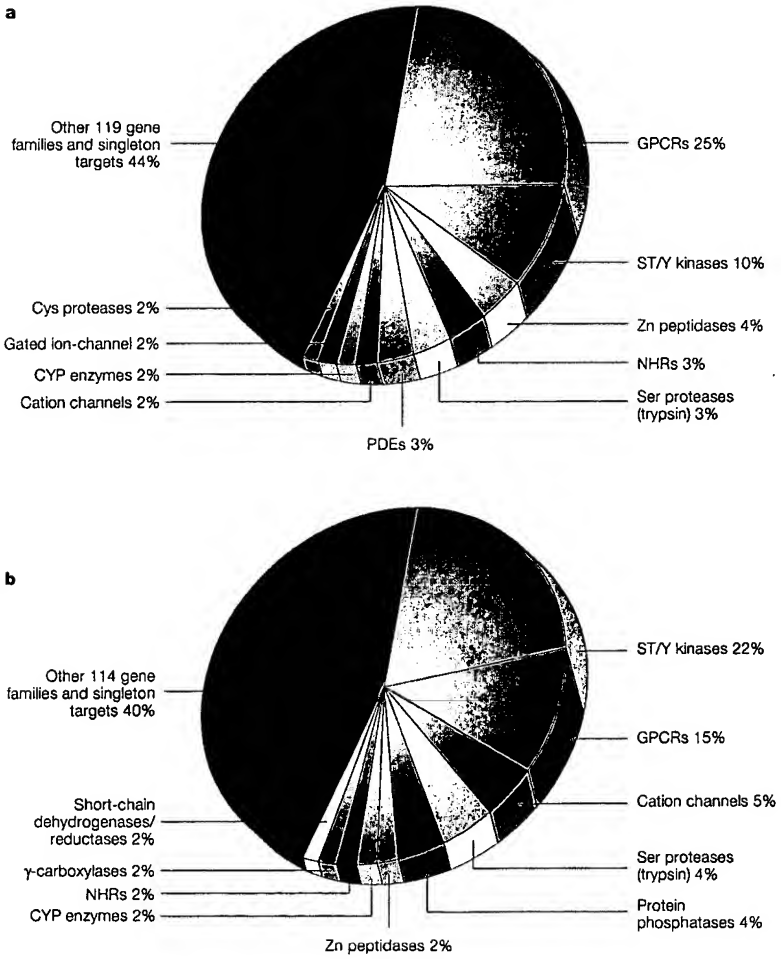


Figure 1 | **Drug-target families.** Gene-family distribution of **a** | the molecular targets of current rule-of-five-compliant experimental and marketed drugs, and **b** | the druggable genome. Serine (Ser)/threonine and tyrosine protein kinases are grouped as one gene family (ST/Y kinases), as are class 1 and class 2 G-protein-coupled receptors (GPCRs). CYP, cytochrome P450; Cys, cysteine; NHR, nuclear hormone receptor; PDE, phosphodiesterase; Zn, zinc.

Expanding the druggable genome

At present, approximately half of the proteins expressed by the genome are functionally unclassified, and of course, some of these might prove to be druggable. However, it is clear from the distribution of the gene-family populations that there are no undiscovered large protein families, which indicates that remaining targets will be members of very small families. Clearly, the number of potential protein targets could be larger than the number of genes, owing to post-translational modifications and assembly of functional complexes; however, this is not likely to increase the number of specific drug-binding sites.

Further evidence for this can be drawn from the observation that despite numerous screening attempts, many targets have failed to show any evidence of binding compounds that are potent and 'drug-like'. This might be a function of the chemical diversity of corporate compound files. However, if druggability is an inherent property of the protein, then *a priori* assessment criteria of potential targets to assess the likelihood of developing a drug against a particular site can be developed. As most drugs bind to discrete binding sites, which can be identified readily by structural analysis, it is possible to filter what we term 'beautiful binding sites' from the wealth of protein-structure data that are available in the Protein Data Bank and are expected soon from structural genomics projects.

As most drugs compete against small molecules for binding sites on proteins, the number of these binding sites is probably a function of the size of the metabolome (the total set of small molecules in an organism). One route to target discovery might therefore

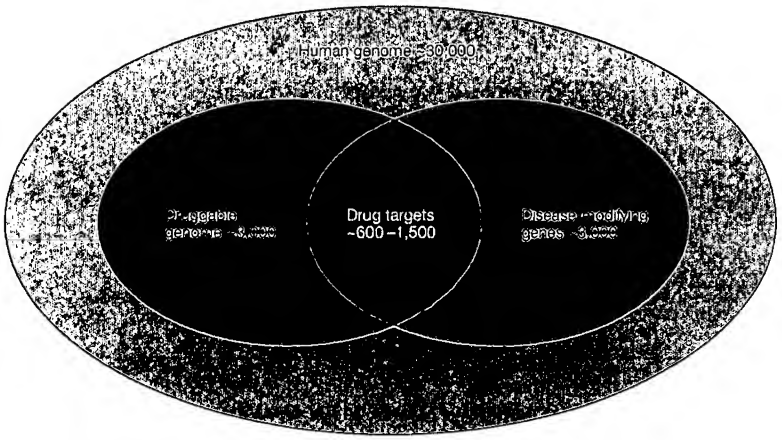


Figure 2 | **Number of drug targets.** The effective number of exploitable drug targets can be determined by the intersection of the number of genes linked to disease and the 'druggable' subset of the human genome.

lie in identifying enzymes and receptors from metabolomic profiling[10–12]. By contrast, the druggability of targets identified by proteomic or transcription-profiling studies is likely to be low.

Druggable does not equal drug target

The ability of a protein to bind a small molecule with the appropriate chemical properties at the required binding affinity might make it druggable, but does not necessarily make it a potential drug target, for that honour belongs only to proteins that are also linked to disease.

Recent estimates propose that there are from 3,000 (REF. 13) to 10,000 (REF. 4) disease-related genes, and large-scale mouse-knockout studies have revealed that only ~10% of all gene knockouts might have the potential to be disease modifying[14], which supports estimates at the lower end of this range.

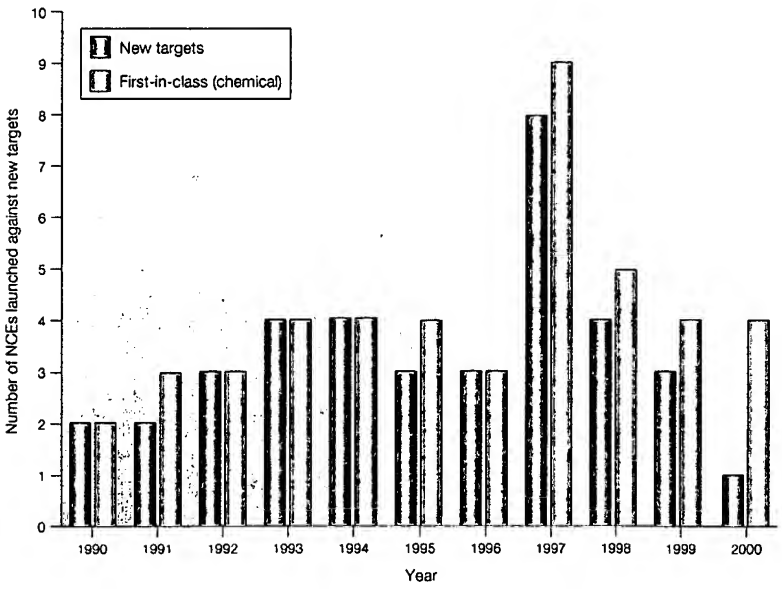


Figure 3 | **Novel drug launches.** The graph shows the number of small-molecule, 'first-in-class' drugs and associated new drug targets that have been launched on the market in the past decade (data derived from collating annual 'This Year's Drugs' reviews of *Drug News & Perspectives*, Prous Science). NCE, new chemical entity.

The potential drug targets that the pharmaceutical industry can exploit are captured in the intersection between the druggable genome and those genes related to disease, as shown in FIG. 2. An analysis of the antifungal targets from the yeast genome indicates that the intersection might be as small as 2–5% of the genome (C.R.G. and J. E. Mills, unpublished observations) — extrapolating to man, this suggests a total of 600–1,500 small-molecule drug targets.

Targets to market

Despite the massive increases in research and development (R&D) investment over the past decade, and the advent of molecular biology, the rate at which drug targets are clinically validated and brought to market is growing rather slowly. On average, new drugs are launched against only four novel targets each year (FIG. 3).

The distribution of target types shown in FIG. 1a is similar to the distribution seen in the original work of Drews[2,3], but to our surprise, of our set of 399 targets with known rule-of-five-compliant agents, we could identify only 120 proteins as the targets of drugs that are actually marketed. This small number of targets calls into question the common assumption that a large number of targets are necessary to build a successful industry[13]. Differentiation between drugs that bind to the same receptor could lead to the development of several distinct classes, targeting a range of diseases.

The overall distribution of launched targets by biochemical class is similar to that observed for all targets with drug-like leads (FIG. 4). Enzymes represent just under half of the launched targets (47%), whereas GPCRs account for 30%. All other classes, such as ion channels and nuclear hormone receptors, account for less than a quarter of the identified launched targets.

Implications

Commercial pressure forces the pharmaceutical industry to focus on developing orally bioavailable small molecules, limiting opportunity to the number of binding sites for such molecules on proteins encoded by the genome. New mechanisms, such as protein drugs, antibody therapies, DNA vaccines and non-oral drug delivery systems, could expand the range of potential targets to those fundamentally not tractable with rule-of-five-compliant therapies.

A comparison of gene-family size with the number of targets in a family that have specific leads shows that many large, druggable gene families are still under-exploited (FIG. 5). The application of high-throughput screening

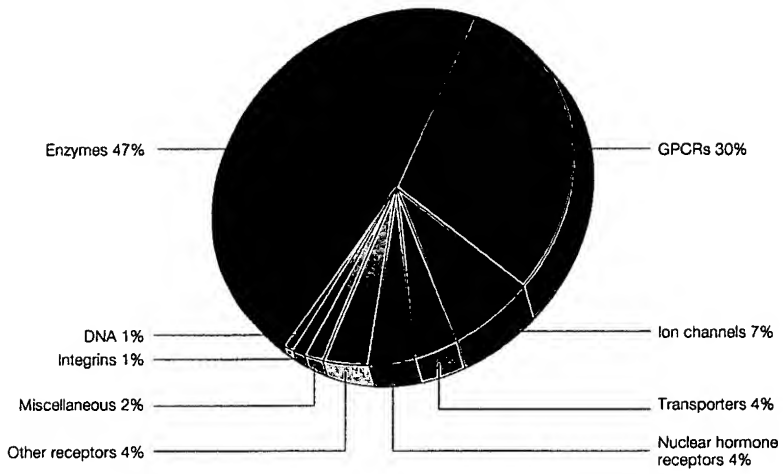


Figure 4 | **Marketed small-molecule drug targets by biochemical class.** GPCR, G-protein-coupled receptor.

in the pharmaceutical industry and the limited number of druggable targets suggest that, within the next decade, the industry could reach a position in which 'hits' or chemical leads are available for most potentially druggable targets. The challenge for the industry will then not necessarily be in the discovery of leads, but in discovering and assessing the therapeutic utility of its leads and druggable targets.

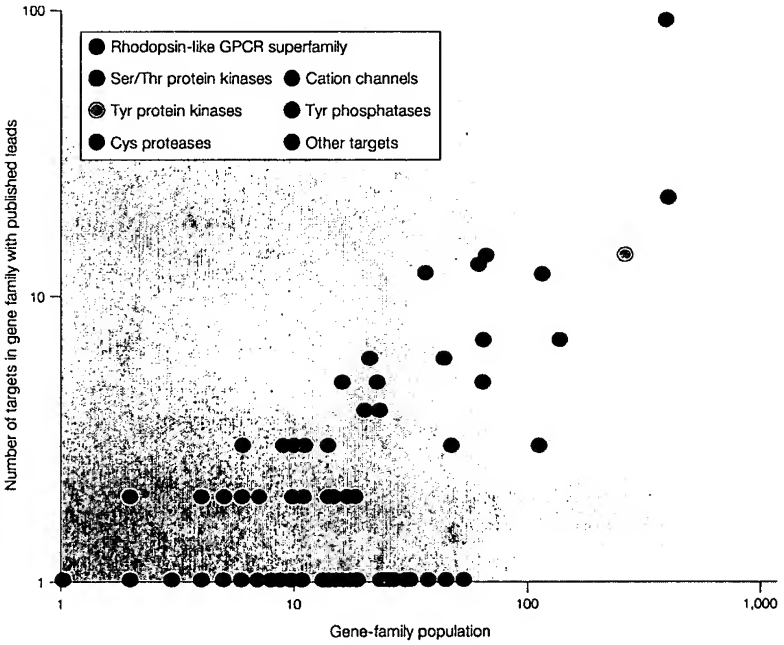


Figure 5 | **Exploitation of the genome, as measured by known leads.** In the past few decades, the pharmaceutical industry has assigned high priority to research into gene families, such as kinases, matrix metalloproteinases and cysteine proteases. Few drugs aimed at these gene families have yet reached the market, although many are progressing through development. Cys, cysteine; GPCR, G-protein-coupled receptor; Ser, serine; Thr, threonine; Tyr, tyrosine.

The limited number of small-molecule drug targets suggests that to exploit the opportunity of the druggable genome in a cost-effective manner, the next round of innovation for the pharmaceutical industry lies not necessarily just in the science, but also in the business models[15].

Andrew L. Hopkins and Colin R. Groom are at the Molecular Informatics, Structure and Design Department, Pfizer Global Research & Development, Sandwich, Kent, CT13 9NJ, UK.
Correspondence to C.R.G.
e-mail: colin_r_groom@sandwich.pfizer.com

doi: 10.1038/nrd892

1. Lipinski, C., Lombardo, F., Dominy, B. & Feeney, P. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).
2. Drews, J. Genomic sciences and the medicine of tomorrow. *Nature Biotechnol.* **14**, 1516–1518 (1996).
3. Drews, J. & Ryser, S. Classic drug targets. *Nature Biotechnol.* **15**, 1318–1319 (1997).
4. Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–1964 (2000).
5. Bailey, D., Zanders, E. & Dean, P. The end of the beginning for genomic medicine. *Nature Biotechnol.* **19**, 207–209 (2001).
6. Apweiler, R. *et al.* The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* **29**, 37–40 (2001).
7. *Proteome Analysis Database* [online], (version analysed with release date 29 Oct 01) <http://www.ebi.ac.uk/proteome/> (2001).
8. Lander, E. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
9. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
10. Tweeddale, H., Notley-McRobb, L. & Ferenci, T. Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool ('metabolome') analysis. *J. Bacteriol.* **180**, 5109–5119 (1998).
11. Fiehn, O. *et al.* Metabolite profiling for plant functional genomics. *Nature Biotechnol.* **18**, 1157–1161 (2000).
12. Raamsdonk, L. *et al.* A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnol.* **19**, 45–50 (2001).
13. Claverie, J.-M. What if there are only 30,000 human genes? *Science* **291**, 1255–1257 (2001).
14. Walke, D. W. *et al.* In vivo drug target discovery: identifying the best targets from the genome. *Curr. Opin. Biotechnol.* **12**, 626–631 (2001)
15. Lehman Brothers. *The Fruits of Genomics* (Lehman Brothers, 2001).
16. Rubin, G. Comparative genomics of the eukaryotes. *Science* **24**, 2204–2215 (2000).

Acknowledgements
We are indebted to J. P. Overington (Inpharmatica, London), A. Alex and L. Beeley for their contributions to the ideas on the physico-chemical limits for protein-binding sites. We also thank R. W. Spencer and C. Lipinski (Pfizer, Groton, Connecticut, USA) for much stimulating discussion.

Online links

FURTHER INFORMATION
InterPro: http://www.ebi.ac.uk/interpro/search.html
Protein Data Bank: http://www.rcsb.org/pdb/
Proteome Analysis Database:
http://www.ebi.ac.uk/proteome/
Access to this interactive links box is free online.

EXHIBIT O

Arthur T. Sands, Current Drug Discovery, Aug. 2002, at 21.

Industrializing breakthrough discovery

Arthur T Sands
Lexicon Genetics, USA

Predicting drug action using mouse knockouts was pioneered by Lexicon; five years later the full potential of gene knockout technology is beginning to be realized. In combination with comprehensive physiological analysis, the technology delivers novel, *in vivo*-validated targets with the potential for the discovery of breakthrough therapeutics.

Innovation in the pharmaceutical industry depends on breakthrough biological discoveries that reveal new targets for therapeutic intervention. These new targets must provide potent new mechanisms of action to block disease by creating favorable alterations in physiology without undesirable side effects. *In vivo* methods of target validation using gene knockouts have revealed truly rare and valuable targets. This fact stands in direct contrast to the popular myth that the human genome contains thousands of viable drug targets. It seems that a simple truth has been all but forgotten in the pursuit of ultra-high-throughput drug discovery: it is the quality of new targets not the quantity that holds promise for replenishing the pharmaceutical industry's drug discovery pipeline.

There is plenty of evidence that more is not necessarily better. The pharmaceutical industry spends $30 billion each year on research and development - three times more than a decade ago - yet the number of new drugs coming to market has not increased. The industry's product innovation bottleneck is especially critical since $38.6 billion in brand name drugs will be coming off patent over the course of the next three years, creating a market void that drug makers are not prepared to fill.

Physiology must guide discovery

To replenish product pipelines, the industry is looking to biotechnology companies to accelerate the identification and validation of new targets. In order to discover which genes among thousands encode breakthrough targets, industry scientists must conduct rigorous physiological assessments to determine which targets to eliminate and which to pursue. Only those targets that demonstrate the potential to maximize therapeutic effects and minimize side effects should be pursued, thereby reducing the failure rate and increasing the overall efficiency of the drug discovery process.

Since a therapeutic alteration in physiology is the desired endpoint of drug discovery, overly reductionist approaches that ignore the complexity of mammalian physiology are inevitably doomed to failure. Computer modeling, DNA microarrays, proteomics and lower model organisms cannot encompass the complexity of mammalian physiology and may actually distract researchers from a more productive pathway to discovery. Even human genetic studies may be problematic, since they are more likely to reveal genes that cause disease rather than drug targets for future cures.

Just as drugs must act within the context of physiology, novel drug targets must be validated within the context of mammalian physiology before precious resources are expended to develop drugs. Grounding genomics in the discipline of physiology can increase success rates, enhance product pipelines and create safer and superior therapeutics, as well as reduce the enormous amount of time and capital expended for the discovery and development of a drug. Those companies who are equipped to rapidly and effectively integrate physiological information into the target selection process will dominate the next generation of successful drug discoveries.

Of knockout mice and men

After a decade of using mouse knockouts, the data on their predictive power in drug discovery is irrefutable. The top 100 selling drugs in 2001 are directed only to 29 drug targets, many with multiple agents addressing the same target. Of these 29 targets, 23 have been knocked out and in every case the knockout mouse was highly predictive as to the on-target effects and side effects of the associated drugs. These observations lay to rest early theoretical concerns regarding the reliability of the mouse knockout technology to recapitulate actions of drugs in mammalian model systems. The recent near completion of the genomic sequence of mouse and man, now available through either public or private DNA sequence databases, has confirmed the high rate of genomic similarity between the two organisms. Indeed, many decades of research have proved the mouse to be an invaluable tool for the evaluation of biological processes relevant to human disease, including immunology, oncology, neurobiology, cardiovascular biology, obesity and many others. Well-established parallels exist between humans and mice on cellular, biochemical and physiological levels.

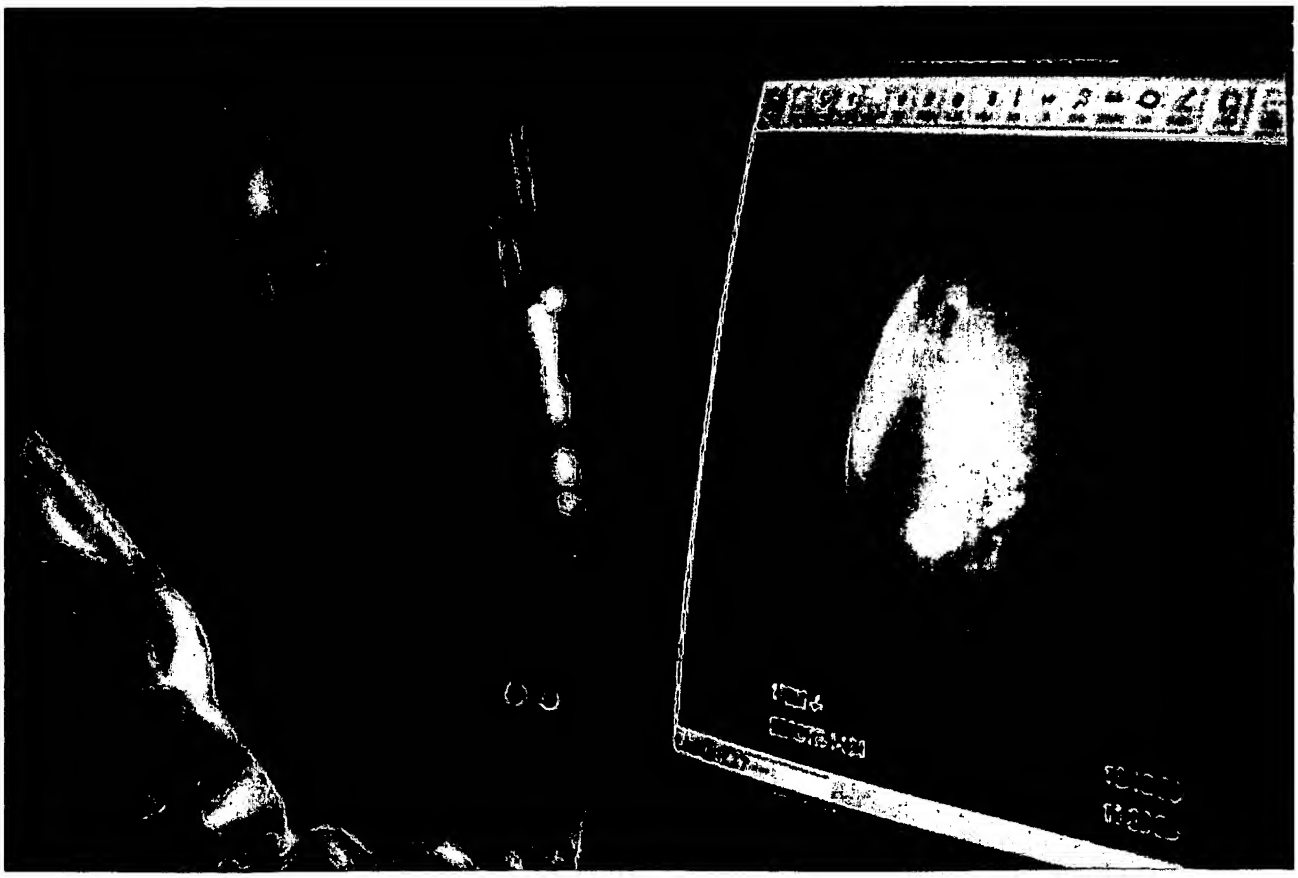
Industrializing discovery

At Lexicon Genetics mouse knockouts are guiding researchers to discover new therapeutic agents which represent the best

physiologic switches in the genome for the treatment of disease. This has required the industrialization of gene targeting, gene trapping and mouse embryonic stem cell technologies, as well as the build-up of significant scientific infrastructure. This infrastructure will allow the company to analyze 5000 genes as mouse knockouts in its Genome5000 program over the next five years. Efforts are concentrated on the unknown function of known gene families for which therapeutic agents can be developed through small molecule chemistry, antibody or therapeutic protein development. These gene families include G protein-coupled receptors, kinases, proteases, ion channels, secreted proteins, transporters and other key enzyme classes. Gene targeting by homologous recombination combined with gene trapping maximizes both selectivity and throughput for large-scale, *in vivo* target validation.

"It seems that a simple truth has been all but forgotten in the pursuit of ultra-high-throughput drug discovery: it is the quality of new targets not the quantity that holds promise for replenishing the pharmaceutical industry's drug discovery pipeline."

The company has deployed a comprehensive, *in vivo* analysis of candidate drug targets that has been modeled after clinical evaluation. Genes analyzed in this way are subject to a superior level of *in vivo* analysis, including physiological function and potential disease indication, providing a robust pipeline of high-value targets. This approach has already proved successful in extracting vital information about the potential medical utility of several new targets in atherosclerosis, diabetes, obesity and CNS disease, among others. Lexicon's physiological analysis utilizes a wide range of the latest medical technologies, including intensive analytical procedures such as the CAT scan for organ system visualization, dual energy X-ray absorptiometry for measurement of percentage fat and lean body mass and bone mineral density, functional magnetic resonance imaging, which allows *in vivo* neurochemical and cardiac analysis, clinical blood and urine chemistries, complete blood cell counts, fluorescent-activated cell sorting, cell-cycle analysis and neurobehavioral testing. Histopathological and gene expression surveys of 55 tissues provide cellular and gene expression data for additional information. Disease challenge models may also be used when indicated to maintain a



Computer modeling, DNA microarrays, proteomics and lower model organisms cannot encompass the complexity of mammalian physiology and may actually distract researchers from a more productive pathway to discovery.

high degree of sensitivity, enabling the detection of subtle phenotypes that may be of significant medical value.

The phenotype derived from the knockout of a specific gene reveals both the potential therapeutic value as well as other target-specific side effects that may be anticipated for a small molecule inhibitor of that target. For instance, a target may display therapeutic potential in inflammation, but might also be critical for renal function. Without a mammalian knockout model, these deleterious target-specific side effects might not be observed until after significant amounts of time and resources have been spent on developing small-molecule compounds and testing them in preclinical or clinical development. When a drug produces a deleterious effect that was not observed in the knockout animal, it suggests that further optimization of the compound's specificity is worthwhile. The ability to produce strong preclinical data to support efficacy and lack of deleterious side effects for a novel target and corresponding lead compound further legitimizes the value of a drug discovery program and provides confidence to move ahead aggressively in development.

Predicting breakthrough therapeutics

Gene knockouts can be viewed as modeling the biological mechanism of drug action by presaging the activity of highly specific antagonists *in vivo*. This information is critical when making decisions regarding target prioritization for a drug discovery enterprise. Since knockout mice have been shown to model drug activity, they provide an unprecedented level of predictive power over the drug discovery process and can be extremely valuable to the pharmaceutical and biotechnology industries. With the effective use of mouse knockout technology, expensive drug discovery activities such as high-throughput screening, medicinal chemistry, preclinical research and clinical trials can be focused on the drug targets that are most likely to lead to breakthrough therapeutics.

Hypothesis-driven gene targeting and gene trapping place physiology and therapeutic potential at the forefront of the drug discovery process and will provide primary data on the physiological function of virtually all members of 'drugable' gene families over the next few years. However, the full power of knockout mouse technology can only be realized when the predictive nature of knockout mouse phenotypes is applied early in the drug discovery process. The combination of mouse gene knockout technology and comprehensive physiological analysis will provide the pharmaceutical industry with novel, *in vivo*-validated targets with clear potential for the discovery of breakthrough therapeutics.

> ***"The top 100 selling drugs in 2001 are only directed to some 29 drug targets, many with multiple agents addressing the same target. Of these 29 targets, 23 have been knocked out and in every case the knockout mouse was highly informative as to the on-target effects and side effects of the associated drugs."***

Arthur T Sands MD, PhD
President & CEO
Lexicon Genetics Inc
8800 Technology Forest Place
The Woodlands
TX 77381-1160
USA

Email: asands@lexgen.com
www.lexgen.com

FURTHER READING

Firn A, Griffith (2002) **Big pharma hopes to get by with some help from old friends.** *Financial Times* May 01.

(2002) **Worldwide functional genomics market expected to reach $2 billion in 2007.** *PR Newswire: Front Line Strategic Consulting* April 18.

EXHIBIT P

Elchebly et al., Science 283:1544 (1999)

Treatise, vol. II (Elsevier/North-Holland, Amsterdam, 1978).
3. T. Jacobs, *Plant Cell* **9**, 1021 (1997); H. Kende and J. A. D. Zeevaart, *ibid.*, p. 1197; W. M. Gray and M. Estelle, *Curr. Opin. Biotechnol.* **9**, 196 (1998).
4. R. Cooke and Y. Meyer, *Planta* **152**, 1 (1981).
5. K. Zhang, D. S. Letham, P. C. L. John, *ibid.* **200**, 2 (1996).
6. A. Jacqmard, C. Houssa, G. Bernier, in *Cytokinins. Chemistry, Activity and Function*, D. W. S. Mok and M. C. Mok, Eds. (CRC, Boca Raton, FL, 1994), pp. 197–215.
7. T. Nagata, S. Ishida, S. Hasezawa, Y. Takahashi, *Int. J. Dev. Biol.* **38**, 321 (1994).
8. M. W. Bayliss, in *The Cell Division Cycle in Plants*, J. A. Bryant and D. Francis, Eds. (Cambridge Univ. Press, Cambridge, 1985), pp. 157–177.
9. C. J. Sherr, *Cell* **73**, 1059 (1993); *Trends Biochem Sci.* **20**, 187 (1995)
10. R. Soni, J. Carmichael, Z. Shah, J. A. H. Murray, *Plant Cell* **7**, 85 (1995).
11. M. Dahl *et al.*, *ibid.*, p. 1847.
12. M. May and C. Leaver, *Plant Physiol.* **103**, 621 (1993).
13. R. A. U. A. Fuerst, R. Soni, J. A. H. Murray, K. Lindsey, *ibid.* **112**, 1023 (1996).
14. *Arabidopsis thaliana* Landsberg *erecta* suspension culture was maintained by weekly 20-fold dilution (*12*, *13*). Early stationary phases (day 7 after previous subculture) were filtered and washed five times with 200 ml of Murashige and Skoog (MS) medium (containing 3% sucrose, but lacking auxin and cytokinin) and maintained in this medium for 24 hours before addition of inducing agents at time $T = 0$ and sampling at times indicated. Identical results were obtained with exponential (day 3) cells similarly treated. Chx (100 μM) was added 1 hour before other inducers and effectively prevents protein synthesis in *Arabidopsis* [S. Abel, M. D. Nguyen, A. Theologis, *J. Mol. Biol.* **251**, 533 (1995); (*22*)].
15. A. Chin-Atkins, S. Craig, C. H. Hocart, E. S. Dennis, A. M. Chaudhury, *Planta* **198**, 549 (1996).
16. Transgenic *Arabidopsis* seedlings carrying the CycD3 promoter linked to a *GUS* reporter gene treated with zeatin also showed twofold to sevenfold induction of GUS but no change in its spatial distribution.
17. Longitudinal sections (8 μm) of 8-week-old vegetative *A. thaliana* Columbia grown under short days, hybridized as described [G. Segers *et al.*, *Plant J.* **10**, 601 (1996)] to 800-nucleotide antisense or full-length sense CycD3 ^{35}S-uridine triphosphate–labeled riboprobes (Promega), were viewed in dark field (autoradiographic signal) or fluorescence (calcofluor staining). For cytokinin induction, 50 μM BAP (Sigma) was applied daily to the apical bud for five successive days.
18. S. E. Clark, R. W. Williams, E. M. Meyerowitz, *Cell* **89**, 575 (1997).
19. Calli were induced by *Agrobacterium*-mediated transformation of *Arabidopsis* roots as described (*20*) with a construct expressing CycD3 under the CaMV 35S promoter. Calli expressed high levels of CycD3 by Northern (RNA) blot (*22*) but could not be regenerated.
20. N. J. Kilby, G. J. Davies, M. R. Snaith, J. A. H. Murray, *Plant J.* **8**, 637 (1995).
21. The CycD3 cDNA without 5′ untranslated region (UTR) was cloned in conditional vector pFLP-SWITCH (G. D. Davies, N. J. Kilby, C. Riou-Khamlichi, J. A. H. Murray, in preparation) and crossed to an hsp-FLP line (*20*). F1 embryos were heat shocked 3 days after pollination, and F2 seeds from individual F1 plants were analyzed for reduced *GUS* transmission. F2 plants from such parents were tested for transmission of activated CycD3 expression and absence of hsp-FLP (see supplementary material available at www.sciencemag.org/feature/data/985481.shl).
22. C. Riou-Khamlichi, A. Jacqmard, J. A. H. Murray, data not shown.
23. Explants cut from rosette leaves of 3-week-old F3 CycD3-expressing and wild-type plants were placed on MS medium with 0.45 μM 2,4-dichlorophenoxyacetic acid (2,4-D) or 2,4-D and 0.44 μM BAP (control). Calli were propagated and trimmed every 10 days; trimming was unnecessary for wild-type calli on medium lacking cytokinin.
24. T. Kakimoto, *Science* **274**, 982 (1996).
25. D. A. Sorrell, B. Combettes, N. Chaubet-Gigot, C. Gigot, J. A. H. Murray, *Plant Physiol.* **119**, 343 (1999).
26. J. C. Mader and D. E. Hanke, *Plant Growth Regul.* **15**, 95 (1996); T. L. Wang, N. P. Everett, A. R. Gould, H. E. Street, *Protoplasma* **106**, 23 (1981).
27. P. Redig, O. Shaul, D. Inzé, M. van Montagu, H. van Onckelen, *FEBS Lett.* **391**, 175 (1996); F. Laureys *et al.*, *ibid.* **426**, 29 (1998).
28. RNA blots on GeneScreen Plus (NEN) were stained with methylene blue to confirm equal loading, hybridized at 42°C, and washed (CycD3 blots: 0.1% standard saline citrate and 0.1% SDS at 42°C for 20 min and at 65°C for 10 min; *Arabidopsis* histone H4 blots: 42°C for 10 min and 55°C for 10 min).
29. Isoprenoid cytokinin determinations by enzyme-linked immunosorbent assay (ELISA) were carried out as described [H. Kraigher, A. Grayling, T. L. Wang, D. E. Hanke, *Phytochemistry* **30**, 2249 (1991)] with SepPak purified extracts of 10-day-old seedlings grown in liquid MS medium. Five ELISA determinations were carried out for at least two separate extractions.
30. T. Kapros *et al.*, *Plant Physiol.* **98**, 621 (1992); J. P. Reichheld, S. Sonobe, B. Clément, N. Chaubet, C. Gigot, *Plant J.* **7**, 245 (1995).
31. Cells were rendered quiescent by incubation of day 7 cells in sucrose-free medium (with hormones) for 24 hours, followed by readdition of 3% sucrose at $T = 0$ and sampling at times indicated (C. Riou-Khamlichi and J. A. H. Murray, in preparation). Incorporation of [^{3}H]thymidine into acid-insoluble material was measured in triplicate relative to background as described (*10*).
32. We thank colleagues for helpful suggestions; N. Kilby, G. Davies, and A. Sessions for advice on the FLP activation system; L. Dehon for in situ hybridizations; D. Hanke for advice and assistance with cytokinin assays; I. Furner for *Arabidopsis* mutants; and A. Inskip for technical assistance. Support of Biotechnology and Biological Sciences Research Council grant P05114 and Pôles d'attraction interuniversitaires belges (Service du Premier Ministre, Services fédéraux des Affaires scientifiques, techniques et culturelles) P4/15 is acknowledged.

24 September 1998; accepted 29 January 1999

Increased Insulin Sensitivity and Obesity Resistance in Mice Lacking the Protein Tyrosine Phosphatase-1B Gene

Mounib Elchebly,[1] Paul Payette,[2] Eva Michaliszyn,[1] Wanda Cromlish,[2] Susan Collins,[2] Ailsa Lee Loy,[1] Denis Normandin,[2] Alan Cheng,[1] Jean Himms-Hagen,[3] Chi-Chung Chan,[2] Chidambaram Ramachandran,[2] Michael J. Gresser,[2] Michel L. Tremblay,[1] Brian P. Kennedy[2]*

Protein tyrosine phosphatase–1B (PTP-1B) has been implicated in the negative regulation of insulin signaling. Disruption of the mouse homolog of the gene encoding PTP-1B yielded healthy mice that, in the fed state, had blood glucose concentrations that were slightly lower and concentrations of circulating insulin that were one-half those of their PTP-1B$^{+/+}$ littermates. The enhanced insulin sensitivity of the PTP-1B$^{-/-}$ mice was also evident in glucose and insulin tolerance tests. The PTP-1B$^{-/-}$ mice showed increased phosphorylation of the insulin receptor in liver and muscle tissue after insulin injection in comparison to PTP-1B$^{+/+}$ mice. On a high-fat diet, the PTP-1B$^{-/-}$ and PTP-1B$^{+/-}$ mice were resistant to weight gain and remained insulin sensitive, whereas the PTP-1B$^{+/+}$ mice rapidly gained weight and became insulin resistant. These results demonstrate that PTP-1B has a major role in modulating both insulin sensitivity and fuel metabolism, thereby establishing it as a potential therapeutic target in the treatment of type 2 diabetes and obesity.

PTP-1B is implicated in the attenuation of the insulin signal (*1*). Mice deficient in the heterotrimeric GTP–binding protein subunit $G_{i_\alpha 2}$ exhibit a phenotype of insulin resistance characteristic of type 2 diabetes that correlates with the increased expression of PTP-1B (*2*). PTP-1B directly interacts with the activated insulin receptor (*3*), and vanadate, a potent nonselective PTP inhibitor, can function as an insulin mimetic both in vitro and in vivo (*4*). However, PTPs other than PTP-1B can also dephosphorylate the activated insulin receptor (*5*). To clarify the role of PTP-1B in insulin action, we generated mice in which the mouse homolog of PTP-1B was disrupted.

The murine gene encoding PTP-1B was cloned from a 129/Sv mouse genomic library and shown to consist of at least nine exons

[1]Department of Biochemistry, McGill University, 3655 Drummond Street, Montreal, Quebec, Canada, H3G 1Y6. [2]Department of Biochemistry and Molecular Biology, Merck Frosst Center for Therapeutic Research, Post Office Box 1005, Pointe Claire-Dorval, Quebec, Canada, H9R 4P8. [3]Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa, Ontario, Canada, K1H 8M5.

*To whom correspondence should be addressed. E-mail: brian_kennedy@merck.com

spanning more than 20 kb (*6*). A targeting vector was designed to delete a portion of the gene that included exon 5 and the tyrosine phosphatase active site in exon 6 and to replace it with the neomycin resistance gene. Two separate embryonic stem cell clones that had undergone homologous recombination and possessed a single integration event were used to microinject Balb/c blastocysts. Chimeric males were mated with wild-type Balb/c females, and heterozygotes from this cross were mated to produce animals homozygous for the PTP-1B mutation (Fig. 1A). Immunoblot analysis of liver microsomes revealed that PTP-1B protein was absent in PTP-1B null mice, and heterozygotes expressed about half the amount of PTP-1B as did wild-type mice (Fig. 1B). PTP-$1B^{-/-}$, heterozygous, and wild-type littermates were born with the same appearance and with the expected mendelian ratio of 1:2:1. PTP-$1B^{-/-}$ mice grew normally, did not show any significant difference in weight gain as compared to wild-type mice, have lived longer than 1.5 years without any sign of abnormality, and are fertile. Complete necropsies were done on male and female wild-type, heterozygous, and homozygous PTP-1B mutant mice 7 to 8 weeks old, and no gross or histological (brain, liver, muscle, pancreas, and testes) differences were observed.

If the role of PTP-1B in the insulin signaling pathway is to dephosphorylate the activated insulin receptor, then mice deficient in PTP-1B might have a sustained insulin response because the insulin receptor would remain phosphorylated and hence be activated longer than in PTP-$1B^{+/+}$ mice. We measured glucose and insulin concentrations in fasted and fed animals (*7*) (Fig. 2). In the fed state, the PTP-$1B^{-/-}$ mice had a significant 13% reduction in blood glucose concentrations, whereas the heterozygotes had an 8% reduction when compared to wild-type mice (Fig. 2A). The PTP-$1B^{-/-}$ mice had circulating insulin concentrations that were about half those of control fed animals (Fig. 2B). Thus, PTP-1B–deficient mice appeared to be more insulin sensitive, because they maintained lower glucose concentrations with significantly reduced amounts of insulin. In the fasted state, there were no significant differences in concentrations of glucose or insulin.

We examined insulin sensitivity in PTP-$1B^{-/-}$, PTP-$1B^{+/-}$, and PTP-$1B^{+/+}$ mice with oral glucose and intraperitoneal insulin tolerance tests (*8*). Administration of a bolus of glucose to PTP-$1B^{-/-}$ mice resulted in a more rapid clearance of glucose than was observed in wild-type mice (Fig. 3A). There was a more pronounced hyperglycemia in the wild-type animals at all time points after glucose administration than in PTP-1B–deficient mice. Increased insulin sensitivity was also observed upon injection of insulin (Fig. 3B). Hypoglycemia was evident 30 and 60 min after injection, but by 120 min after injection glucose concentrations were returning to normal values in wild-type mice, whereas the PTP-1B–deficient mice remained hypoglycemic ($P < 0.02$). The PTP-

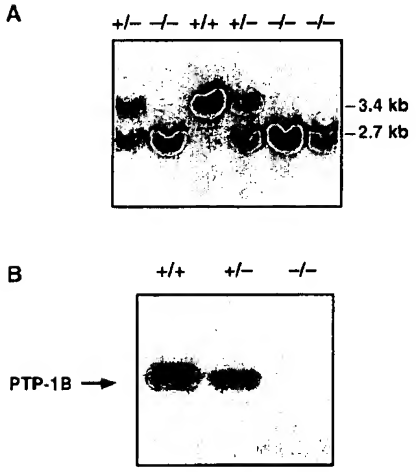


Fig. 1. Gene targeting of the PTP-1B locus. (**A**) Representative genomic Southern blot analysis of tail DNA digested with Bam HI from a heterozygous cross resulting in wild-type (+/+), heterozygous (+/−), and homozygous (−/−) PTP-1B offspring. (**B**) PTP-1B immunoblot analysis of liver membrane samples from PTP-$1B^{+/+}$, PTP-$1B^{+/-}$, and $PTP^{-/-}$ mice.

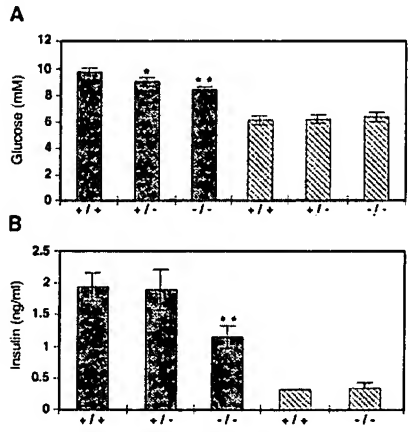


Fig. 2. Serum concentrations of glucose and insulin in animals fed ad libitum or fasted overnight. (**A**) Glucose and (**B**) insulin concentrations were determined as described (*7*). Dark bars indicate fed group [(A) and (B), $n = 19$ to 21]; light bars indicate fasted group [(A), $n = 8$ to 10; (B), $n = 6$]. Values are given as means ± SEM. Statistical analysis was done with a two-tailed, unpaired, Student's t test. Compared to the wild type, $*P = 0.06$ and $**P \leq 0.01$.

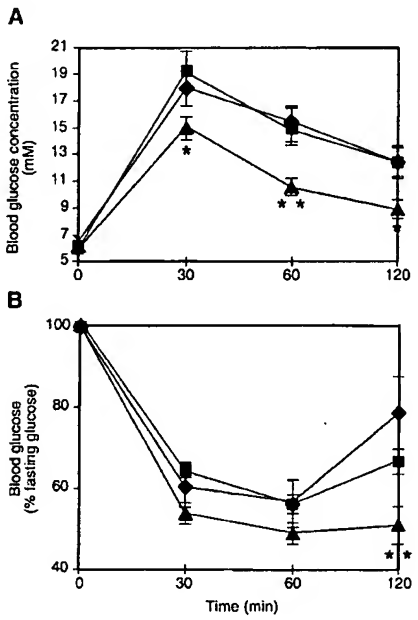


Fig. 3. Glucose and insulin tolerance tests in PTP-$1B^{+/+}$ (diamonds), PTP-$1B^{+/-}$ (squares), and PTP-$1B^{-/-}$(triangles) mice. (**A**) Glucose tolerance of male mice 10 to 14 weeks old ($n = 11$). (**B**) Insulin tolerance of male mice 10 to 14 weeks old ($n = 5$ to 6). Data are presented as means ± SEM. Statistical analysis was done with a two-tailed, unpaired, Student's t test. Compared to the wild type, $*P < 0.05$ and $**P < 0.02$.

Table 1. Fasting glucose, triglyceride, and insulin levels of male wild-type, heterozygote, and PT-$1B^{-/-}$ littermates fed a normal or high-fat diet. Values are given as the means ± SEM. Statistical analysis was done with a two-tailed, unpaired, Student's t test. ND, not determined.

Diet	PTP-$1B^{+/+}$		PTP-$1B^{+/-}$		PTP-$1B^{-/-}$	
	Normal	High fat	Normal	High fat	Normal	High fat
Glucose (mM)	6.1 ± 0.3	8.1 ± 0.6*	6.2 ± 0.3	7.3 ± 0.6	6.3 ± 0.3	7.0 ± 0.4
Triglycerides (mM)	1.84 ± 0.76	2.41 ± 0.19	1.43 ± 0.44	2.44 ± 0.32	0.86 ± 0.18*	1.46 ± 0.15*
Insulin (ng/ml)	0.30 ± 0.02	0.98 ± 0.32	ND	0.97 ± 0.30	0.33 ± 0.08	0.45 ± 0.14*

*$P < 0.05$ ($n = 6$ to 10).

1B$^{+/-}$ mice did not show altered glucose tolerance as compared to that in wild-type mice, although glucose concentration at 120 min after injection of insulin appeared to be intermediate between that of wild-type and PTP-1B$^{-/-}$ mice (Fig. 3).

Binding of insulin to its receptor results in autophosphorylation of the receptor on tyrosines 1146, 1150, and 1151 in the kinase regulatory domain (*9*). This causes activation of the insulin receptor tyrosine kinase, which phosphorylates the various insulin receptor substrate (IRS) proteins that propagate the insulin signaling event (*9*). If PTP-1B dephosphorylates the activated insulin receptor, then the increased insulin sensitivity observed in the PTP-1B$^{-/-}$ mice may be due to increased or prolonged phosphorylation of the receptor (or both). We therefore measured tyrosine phosphorylation of the insulin receptor in liver and muscle tissue (Fig. 4, A and B) after exposure to insulin (*10*). The kinetics of insulin receptor phosphorylation in the liver were significantly different in PTP-1B$^{-/-}$ mice than those observed in the PTP-1B$^{+/+}$ mice. The amount of insulin receptor phosphorylation was the same for both PTP-1B$^{-/-}$ and PTP-1B$^{+/+}$ animals 1 min after injection. However, by 5 min after injection, phosphorylation decreased to about 50% of maximal in the PTP-1B$^{+/+}$ animals, whereas no reduction was observed in the PTP-1B$^{-/-}$ mice. A greater effect on insulin receptor phosphorylation was observed in muscle of the PTP-1B$^{-/-}$ mice. The absolute amount of receptor phosphorylation was increased about twofold in muscle from PTP-1B$^{-/-}$ mice as compared to that in muscle from PTP-1B$^{+/+}$ animals ($P < 0.05$) (Fig. 4B). The amount of insulin receptor phosphorylation in muscle did not significantly change between 2 to 6 min after insulin injection in either PTP-1B$^{-/-}$ or PTP-1B$^{+/+}$ samples. No difference in the amount of insulin receptor expression was detected between PTP-1B$^{+/+}$ and PTP-1B$^{-/-}$ mice, as determined by protein immunoblotting (*11*).

To confirm that the increased phosphorylation of the insulin receptor in the muscle of insulin-treated PTP-1B$^{-/-}$ mice reflects increased kinase activity, phosphorylation of IRS-1 was also measured (*12*) (Fig. 4C). Phosphorylation of IRS-1 was increased in insulin-treated muscle from PTP-1B$^{-/-}$ mice in comparison to muscle from wild-type animals ($P < 0.05$). We have also examined phosphorylation of another receptor tyrosine kinase, the epidermal growth factor receptor, and found no difference between PTP-1B wild-type and deficient animals (*11*). The increased insulin receptor phosphorylation in muscle and its sustained phosphorylation in the liver probably accounts for the enhanced insulin sensitivity observed in the PTP-1B$^{-/-}$ mice.

The disruption of the PTP-1B gene demonstrates that altering the activity of PTP-1B can modulate insulin signaling in vivo. To determine the effect of the loss of PTP-1B activity on insulin resistance, PTP-1B–deficient, wild-type, and heterozygote littermates were subjected to a diet high in fat (50% of calories from fat) and calories (5286 kcal kg^{-1}; Bioserve, NJ) (*13*) that normally results in obesity-induced insulin resistance (*14*). During the 10 weeks the mice were on this diet, male and female wild-type littermates rapidly gained weight, whereas PTP-1B$^{-/-}$ and PTP-1B$^{+/-}$ mice were substantially protected from diet-induced weight gain (Fig. 5). The amount of food consumed by the animals did not differ, which indicates that decreased expression of PTP-1B (heterozygotes have about half the amount of PTP-1B as that in wild-type animals) can influence dietary-induced obesity.

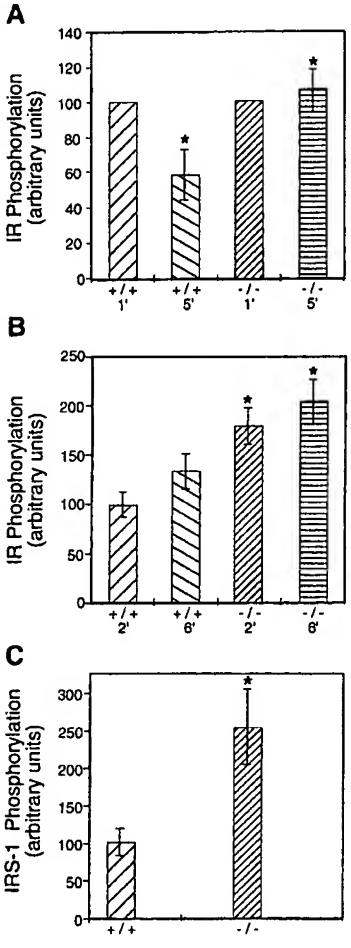


Fig. 4. Increased and prolonged tyrosine phosphorylation of the insulin receptor in PTP-1B$^{-/-}$ mice. (**A**) Time course of tyrosine phosphorylation of the insulin receptor (IR) β subunit in the liver after insulin treatment in PTP-1B$^{+/+}$ and PTP-1B$^{-/-}$ mice. Quantification of the insulin receptor β subunit phosphotyrosine content from immunoblots was performed by densitometry. Data are presented by setting the amount of phosphorylation at 1 min to 100 and the subsequent amount after 5 min for the same animal relative to this value. The results are from five PTP-1B$^{-/-}$ and PTP-1B$^{+/+}$ mice each, from three separate experiments. (**B**) Phosphorylation of the insulin receptor β subunit in muscle of insulin-treated PTP-1B$^{+/+}$ and PTP-1B$^{-/-}$ mice. The quantified data from immunoblots ($n = 6$, from two separate experiments) are presented as arbitrary units. (**C**) Insulin-stimulated tyrosine phosphorylation of IRS-1 from muscle of PTP-1B$^{+/+}$ and PTP-1B$^{-/-}$ mice 2 min after injection ($n = 3$). The quantified data are presented as arbitrary units. Data are means ± SEM. Statistical analysis was done with a two-tailed, unpaired, Student's *t* test comparing in (A) the 5-min to the 1-min time point value and in (B) the PTP-1B$^{-/-}$ mice 2- and 6-min time point values to the respective values of the PTP-1B$^{+/+}$ mice ($*P < 0.05$).

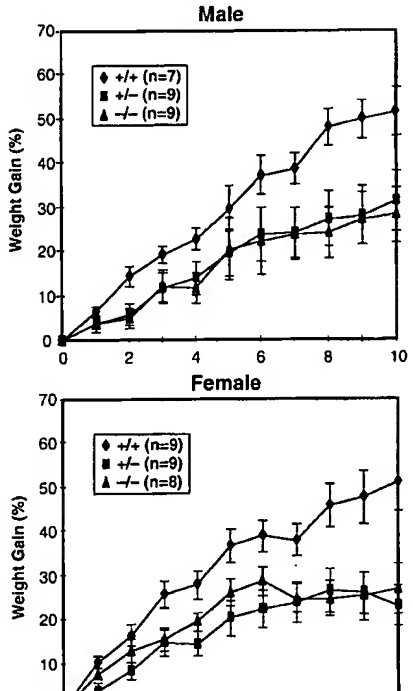


Fig. 5. Resistance of PTP-1B null and heterozygous mice to diet-induced obesity. The percent weight gain of male and female wild-type (diamonds), heterozygous (squares), and homozygous (triangles) littermates fed a high-fat diet for 10 weeks is shown. The starting weight (male: +/+, 27.6 ± 1.4 g; +/–, 28.5 ± 1.2 g; and –/–, 26.3 ± 1.2 g; female: +/+, 22.1 ± 0.8 g; +/–, 22.2 ± 0.8 g; and –/–, 21.5 ± 0.8 g) and final weight (male: +/+, 41.4 ± 1.3 g; +/–, 37.2 ± 2.0 g; and –/–, 33.5 ± 1.6 g; female: +/+, 33.3 ± 1.7 g; +/–, 27.3 ± 1.3 g; and –/–, 27.2 ± 1.4 g) of animals put on the high-fat diet are indicated. The final weight was significantly different ($P < 0.05$) for PTP-1B null and heterozygous mice as compared to wild-type mice, except for male wild-type mice versus heterozygotes ($P = 0.1$).

We examined the effect of the high-fat diet on insulin sensitivity in these animals. Glucose and insulin concentrations in the serum of fasting animals were measured, and glucose and insulin tolerance tests were done on all groups of animals [values for males are presented; females had essentially the same response (11)]. In wild-type mice, the high-fat diet produced a 30% increase ($P < 0.05$) in fasting blood glucose concentrations and a threefold increase in serum insulin concentrations (Table 1) as compared to wild-type mice on a normal diet. In contrast, the PTP-1B$^{-/-}$ animals maintained glucose and insulin concentrations while on the high-fat diet, which were not significantly different from those in animals on a normal diet (Table 1). PTP-1B heterozygotes on a high-fat diet showed increased fasting concentrations of circulating insulin but had fasting glucose concentrations that were not significantly different from those in animals on a normal diet (Table 1). PTP-1B$^{-/-}$ mice also had enhanced insulin sensitivity as compared to their wild-type littermates in both glucose and insulin tolerance tests (Fig. 6, A and B). The difference in insulin sensitivity between the PTP-1B$^{-/-}$ and PTP-1B$^{+/+}$ mice became more augmented on the high-fat diet because of the obesity-induced insulin resistance of the wild-type mice. Obesity-induced insulin resistance results in a reduction in insulin receptor phosphorylation and hence in insulin signaling (*14*). Examination of insulin-stimulated receptor phosphorylation in mice on the high-fat diet revealed that there was a much greater difference in the amount of insulin receptor phosphorylation between the PTP-1B wild-type and deficient animals (Fig. 6C). This increased difference appears to be due to both a reduction in the amount of insulin-stimulated receptor phosphorylation in wild-type mice and a slight increase in the amount of insulin-stimulated receptor phosphorylation in the PTP-1B$^{-/-}$ mice. The PTP-1B heterozygotes on the high-fat diet also appeared to maintain a better response to insulin-stimulated receptor phosphorylation than did wild-type animals (Fig. 6).

The reason for the obesity resistance observed in the PTP-1B$^{-/-}$ mice is unclear at this time, but analysis of triglyceride concentrations indicates that fat metabolism has been affected in these animals. The PTP-1B$^{-/-}$ mice on either diet had significantly lower triglyceride concentrations than did wild-type and heterozygous mice (Table 1). We also examined insulin-stimulated receptor phosphorylation in adipose tissue and found no significant difference between wild-type and PTP-1B-deficient animals on either diet (*11*). Thus, PTP-1B–deficient mice appear to show tissue-specific insulin sensitivity; muscle and liver have an enhanced insulin sensitivity, whereas adipose tissue remains unchanged relative to wild-type mice. The effects that were observed in adipose tissue could be the result of some compensatory mechanism such as the up-regulation of some other PTP [although no difference was observed in the amount of PTP activity between wild-type and PTP-1B$^{-/-}$ tissue extracts (*11*)] or could possibly be due to the overall enhanced insulin sensitivity of the PTP-1B–deficient animal.

The data presented identify PTP-1B as having a major role in the insulin signaling pathway. What this function is remains to be clarified, but the simplest explanation would be that it dephosphorylates the activated insulin receptor. Recently, the disruption of the leukocyte antigen–related (LAR) PTP, which has also been suggested to affect the insulin signaling cascade, has been described. The targeted mutagenesis of LAR produced mice with impaired mammary gland development (*15*) but with blood glucose concentrations within the normal range (*16*). In contrast, the LAR-deficient mice generated by insertional mutagenesis had body weights that were half those of control mice and were insulin resistant (*17*). The reason for the difference in phenotype between these two LAR-deficient mice strains is unknown.

We have shown that the loss of PTP-1B activity causes enhanced insulin sensitivity and resistance to weight gain in mice. These results make PTP-1B a potential therapeutic target for the treatment of type 2 diabetes and obesity.

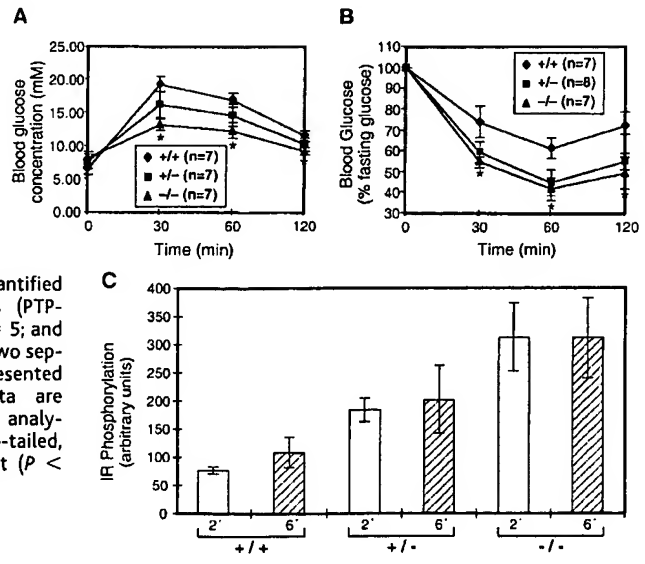


Fig. 6. Insulin tolerance of mice lacking PTP-1B and of wild-type mice on a high-fat diet. (**A**) Glucose and (**B**) insulin tolerance tests of male mice ($n = 7$ to 8) on a high-fat diet. (**C**) Insulin-stimulated insulin receptor tyrosine phosphorylation in muscle of fat-fed mice. The quantified data from immunoblots (PTP-1B$^{-/-}$ and PTP-1B$^{+/+}$, $n = 5$; and PTP-1B$^{+/-}$, $n = 3$; from two separate experiments) are presented as arbitrary units. Data are means ± SEM. Statistical analysis was done with a two-tailed, unpaired, Student's t test ($P < 0.05$).

References and Notes

1. M. F. Cicirelli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5514 (1990); H. Maegawa *et al.*, *J. Biol. Chem.* **270**, 7724 (1995); H. Chen *et al.*, *ibid*, **272**, 8026 (1997); F. Ahmad, P.-M. Li, J. Meyerovitch, B. J. Goldstein, *ibid.*, **270**, 20503 (1995).
2. C. M. Moxham and C. C. Malbon, *Nature* **379**, 840 (1996).
3. B. L. Seely *et al.*, *Diabetes* **45**, 1379 (1996); D. Bandyopadhyay *et al.*, *J. Biol. Chem.* **272**, 1639 (1997).
4. R. L. Khandelwal and S. Pugazhenthi, *Mol. Cell. Biochem.* **153**, 87 (1995).
5. K. K. Jacob, J. Sap, F. M. Stanley, *J. Biol. Chem.* **273**, 4800 (1998); N. Hashimoto, E. P. Feener, W.-R. Zhang, B. J. Goldstein, *ibid.* **267**, 13811 (1992); P. Chiarugi *et al.*, *Biochem. Biophys. Res. Commun.* **238**, 676 (1997); N. P. H. Moller *et al.*, *J. Biol. Chem.* **270**, 23126 (1995).
6. The mouse PTP-1B gene was isolated by screening a Lambda FIX II 129/SvJ mouse genomic library (Stratagene), and the targeting vector was designed to delete exon 5 and the active site Cys^{215} in exon 6 of the gene. The vector consisted of a 5.5-kb Hind III–Eco RI mouse PTP-1B genomic fragment upstream of exon 5, the neomycin resistance gene driven by the phosphoglycerate kinase (PGK) promoter, a 1.6-kb Xba I–Xho I mouse PTP-1B genomic fragment downstream of exon 6, and the herpes simplex virus–thymidine kinase gene driven by the PGK promoter. Electroporation of the linearized vector into 129/Sv embryonic stem cell line J1 [E. Li, T. H. Bestor, R. Jaenisch, *Cell* **69**, 915 (1992)] and selection of G418 (400 μg/ml)–resistant transformants were done as described [K. E. You-Ten *et al.*, *J. Exp. Med.* **186**, 683 (1997)]. Colonies resistant to G418 were analyzed for homologous recombination by Bam HI digestion of genomic DNA followed by Southern (DNA) blotting and hybridization with a 3′ probe from outside the recombination region. Generation of chimeric and mutant mice was done as described [K. E. You-Ten *et al.*, *J. Exp. Med.* **186**, 683 (1997)]. Genotyping was done by Southern blotting. Immunoblot analysis of PTP-1B expression in liver membranes (25 μg per lane) of PTP-1B$^{+/+}$, PTP-1B$^{+/-}$, and PTP$^{-/-}$ mice was done with a rabbit polyclonal antibody to a PTP-1B NH_2-terminal peptide detected with enhanced chemiluminescence (NEN).
7. Male mice were used. Blood was collected from the orbital sinus of anesthetized mice, and serum was prepared. Serum glucose and triglyceride concentrations were determined with a Vitros analyzer, and a radioimmunoassay (Linco, St. Charles, MO) was used to measure insulin concentrations.
8. Glucose tolerance was performed after an overnight fast by oral administration of glucose [1 g per kilogram of body weight (1 g/kg)]. Insulin tolerance tests were performed after an overnight fast by intraperitoneal injection of 0.75 U/kg of human insulin (1 U ~ 40 μg) (Lilly). Blood was withdrawn from the tail, and glucose concentrations were determined with a One Touch Basic glucometer (Lifescan Canada Ltd., Burnaby, Canada).

9. M. F. White and C. R. Kahn, *J. Biol. Chem.* **269**, 1 (1994).
10. After an overnight fast, mice were anesthetized, the abdominal cavity was exposed, and human insulin (5 U) (Lilly) or saline was injected as a bolus into the inferior vena cava [E. Araki *et al.*, *Nature* **372**, 186 (1994); P. L. Rothenberg *et al.*, *J. Biol. Chem.* **266**, 8302 (1991)]. One minute after injection, a small piece of liver was excised and immediately frozen in liquid nitrogen. This was followed at 2 and 3 min for quadriceps muscle and abdominal fat, respectively, and the same procedure was again repeated at 5, 6, and 7 min. Immunoprecipitation of the insulin receptor β subunit was done as follows: the tissue (liver, muscle, or fat) was homogenized on ice in 50 mM tris (pH 7.5); 150 mM NaCl; 1 mM pyrophosphate; 100 μM pervanadate (a potent PTP inhibitor) [G. Huyer *et al.*, *J. Biol. Chem.* **272**, 843 (1997)]; and a protease inhibitor cocktail tablet, Complete (Boehringer Mannheim). A membrane fraction was prepared by centrifugation at 100,000*g* for 1 hour, and the protein concentration was determined. Membrane protein (200 μg from liver or muscle, 100 μg from fat) was solubilized in immunoprecipitation buffer (RIPA) [150 mM NaCl, 10 mM phosphate buffer (pH 7.5), 1% NP-40, 1% Na deoxycholate, and 0.1% SDS], and immunoprecipitation of the insulin receptor β subunit was done overnight at 4°C with the use of antibody to the insulin receptor (1 μg) (C-19, Santa Cruz Biotechnology, Santa Cruz, CA) followed by a 60-min incubation with 50 μl of a 50% slurry of protein G–Sepharose (Pharmacia Biotech). The sample was washed three times in 1 ml of RIPA buffer, and samples were resolved by SDS–polyacrylamide gel electrophoresis (8% gel). The gel was then transferred onto a polyvinyl difluoride membrane, and the phosphotyrosine was detected with antibody to phosphotyrosine (4G10) coupled to horseradish peroxidase (Upstate Biotech, Waltham, MA) and developed with enhanced chemiluminescence (NEN). The same blot was washed in 62.5 mM tris (pH 6.7), 2% w/v SDS, and100 mM β-mercaptoethanol for 30 min at 55°C; rinsed; reprobed with the antibody to the insulin receptor β subunit (C-19, Santa Cruz Biotechnology), followed by a second antibody to rabbit immunoglobulin G coupled to horseradish peroxidase (Amersham); and detected by enhanced chemiluminescence (NEN). The phosphotyrosine and β-subunit signals from the x-ray films of the exposed blots were quantified by densitometry (Molecular Dynamics), and amounts of phosphotyrosine were normalized to the amount of β subunit present in each sample.
11. P. Payette, M. Elchebly, M. L. Tremblay, B. P. Kennedy, unpublished data.
12. Immunoprecipitation of IRS-1 from the cytosolic fraction of muscle from insulin-treated mice was done with two rabbit polyclonal antibodies to IRS-1 (C-20, which binds to the COOH-terminus and A-19, which binds to the NH_2-terminus).
13. Male and female PTP-1B$^{-/-}$ mice and their wild-type and heterozygous littermates at 7 to 8 weeks of age were fed a high-fat high-carbohydrate diet ad libidum (Diet F3282, Bioserve, NJ) and monitored for the following 10 weeks. Body weight was measured weekly, and food consumption was closely monitored. After 10 weeks, concentrations of glucose, insulin, and triglycerides in the serum of fasted animals were measured, and glucose and insulin tolerance tests were done.
14. K. T. Uysal, S. M. Wiesbrock, M. W. Marino, G. S. Hotamisligil, *Nature* **389**, 610 (1997).
15. R. Q. J. Schaapveld *et al.*, *Dev. Biol.* **188**, 134 (1997).
16. K. Norris *et al.*, *FEBS Lett.* **415**, 243 (1997).
17. J.-M. Ren *et al.*, *Diabetes* **47**, 493 (1998).
18. We thank D. Moller, J. Mudgett, E. Asante-Appiah, and J. Evans for discussions. M.E. is a fellowship recipient and M.L.T. is a Chercheur-Boursier Senior from the Fonds de la Recherche en Sante du Quebec. Supported in part by a grant from the Medical Research Council (MRC)/Pharmaceutical Manufacturers Association of Canada (M.L.T.) and the MRC (J.H.-H.).

20 August 1998; accepted 29 January 1999

EXHIBIT Q

American Association of Clinical Endocrinologists Medical Guidelines for Clinical Practice for the Diagnosis and Treatment of Hyperandrogenic Disorders, Endocrine Practice 7:121 (2001)

AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS MEDICAL GUIDELINES FOR CLINICAL PRACTICE FOR THE DIAGNOSIS AND TREATMENT OF HYPERANDROGENIC DISORDERS

Hyperandrogenic Disorders Task Force

Chairman

Neil F. Goodman, MD, FACE

Committee Members

Maya B. Bledsoe, MD, FACE
Rhoda H. Cobin, MD, FACE
Walter Futterweit, MD, FACP, FACE
Joseph W. Goldzieher, MD, FACE
Steven M. Petak, MD, JD, FACE
Keith D. Smith, MD, FACE
Emil Steinberger, MD, FACE

Reviewers

Robert J. Anderson, MD, FACE
Donald A. Bergman, MD, FACE
Zachary T. Bloomgarden, MD, FACE
Richard A. Dickey, MD, FACP, FACE
Pasquale J. Palumbo, MD, MACE
Anne L. Peters, MD
Herbert I. Rettinger, MD, FACE
Helena W. Rodbard, MD, FACE
Harvey A. Rubenstein, MD, FACE

AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS MEDICAL GUIDELINES FOR CLINICAL PRACTICE FOR THE DIAGNOSIS AND TREATMENT OF HYPERANDROGENIC DISORDERS

MISSION STATEMENT

The purpose of this document is to provide guidelines for the diagnosis and treatment of hyperandrogenic disorders in women. These may range from simple hirsutism, without clearly demonstrable biochemical features of hyperandrogenism, to frank virilization. The symptoms and signs include those that involve the pilosebaceous unit (PSU)—that is, hirsutism, acne, and alopecia—and those that involve the female reproductive system—amenorrhea and infertility. Hyperandrogenism is also a forerunner of serious cardiovascular problems (for example, hypertension, microvascular disease, and dyslipidemias) and other metabolic disorders (such as type 2 diabetes mellitus). Thus, clinical endocrinologists seem optimally qualified to address these disorders.

These guidelines are not to be considered an extensive review of the literature on this topic or an exhaustive analysis of recent advances in this field. They are intended to be what the word guidelines implies: a brief summary of the accepted scientific information and views on this topic as well as suggestions for the diagnosis and treatment of these disorders. The presented material is in a form that can be easily used and applied to practical clinical situations encountered by endocrinologists.

INTRODUCTION

"Hyperandrogenism" is a term used to describe the most common clinical signs in women with hyperandrogenemia: hirsutism, acne, and alopecia. It is also the hyperandrogenic state that drives the various other pathologic conditions in a wide range of tissues and organ systems. The most commonly diagnosed conditions associated with hyperandrogenism in reproductive-age women are ovulatory disorders and polycystic ovary syndrome (PCOS). PCOS has been estimated to affect 5 to 10% of women in this age-group (1). The manifestations of hyperandrogenism are detected and frequently classified relative to the medical specialty practiced by the attending physician. For example, the dermatologist may note some of the cutaneous manifestations (acne, hirsutism, and alopecia), the gynecologist addresses the menstrual dysfunction (amenorrhea, oligomenorrhea, menorrhagia, and metrorrhagia) and the underlying ovarian dysfunction (anovulation, oligo-ovulation, pelvic pain, ovarian cysts, and infertility), the pediatrician deals with associated congenital adrenal hyperplasia (CAH) or ambiguous genitalia, the internist diagnoses the dyslipidemias, hypertension, and impaired glucose tolerance that may be associated with long-standing hyperandrogenemia, and the endocrinologist usually encounters patients with hyperandrogenism who have symptoms and signs of hirsutism, acne, and insulin resistance.

Regardless of which specialty assumes the responsibility for the management of the patient's condition, the disease is an endocrine disorder—a hyperandrogenic state—that may precipitate pathologic events in various organ systems (for example, PSU activation, ovulatory dysfunction, ovarian cysts, menstrual dysfunction, insulin resistance, infertility, and cardiovascular disorders). The disease is usually caused by excessive androgen production by the ovaries, the adrenal glands, or both. It can also be attributable to an abnormality of the androgen receptor mechanisms in the target cell (the incidence of this abnormality is unknown). Hyperandrogenism caused by abnormalities in peripheral metabolism of steroids is also rare.

These guidelines deal primarily with the following factors:

- Symptoms of hyperandrogenism
- Diagnostic evaluation of hyperandrogenism
- Determination of the site (or sites) of excess androgen production
- Determination of pathologic and pathophysiologic changes at the site of androgen hypersecretion
- Therapy for hyperandrogenism

SYMPTOMS OF HYPERANDROGENISM

The Pilosebaceous Unit

Acne

Acne is commonly present in female adolescents. It occurs in almost 50% of teenage subjects. Persistence of acne into the late teens or 20s, however, should alert the pediatrician or the endocrinologist to the possibility of hyperandrogenism. This scenario is particularly likely if the acne is resistant to the usual dermatologic treatment strategies and is associated with hirsutism or menstrual

dysfunction. Under these circumstances, acne should be considered a sign of hyperandrogenism that necessitates appropriate diagnostic investigation (2,3). The variability in the genetic susceptibility of the PSU to androgen stimulation is associated with variations in the clinical expression of the hyperandrogenism (4). Women with acne alone may have levels of plasma testosterone as high as those with hirsutism, with or without acne. Similarly, no correlation exists between the severity of acne and the plasma free testosterone levels (5). As with other manifestations of the response of the PSU to hyperandrogenism, persistent acne should prompt an investigation of other androgen-related disorders. Of note, ovulatory dysfunction is common in young women with acne (6). In one study of women seen in consultation primarily for acne, 45% of cases were associated with polycystic ovaries (7).

Hirsutism

Hirsutism is defined as excessive hair growth in women, occurring in anatomic areas where the hair follicles are most androgen sensitive. Hirsutism manifests as an excessive or inappropriate development and growth of the PSU (6). Androgens cause transformation of vellus (fine, soft, unpigmented) hair to terminal hair in androgen-dependent areas of hair growth.

The amount, distribution, and progression of human body hair have racial, familial, genetic, and hormonal influences. The massive hairiness of aboriginal Ainu in Japan is considered a sign of racial purity. Extreme genetic (nonracial) hairiness has been documented in art (8). Nordic and Anglo-Saxon European peoples, indigenous natives of North and South America, and African Americans are generally less hairy than are Mediterranean peoples. About one in three non-Scandinavian or non-Asian women has some hair on the upper lip, periareolar, and suprapubic areas. East Asians tend to be less hairy than Euro-Americans, with no difference in testosterone levels (9). Hormonal levels are also normal in prepubertal simple hypertrichosis (10), which is the appropriate term for hairiness of nonhormonal origin, in contrast to hirsutism and its extreme, virilization. Familial hairiness may be genetic or hormonal. Age-dependent variation in hairiness has also been reported (11).

The presence of substantial numbers of terminal hairs on the lower back, sternum, abdomen, shoulders, buttocks, and inner thighs is considered abnormal. The degree and extent of hirsutism are clinically evaluated by using the Ferriman-Gallwey scale (12) or a modification (13,14) (Fig. 1 and Table 1). This method is useful for assessment and follow-up of the response to therapy.

The variability in the inherent sensitivity of the PSU to androgens is demonstrated by the fact that even though plasma testosterone levels may be substantially increased, the patient may not exhibit appreciable hirsutism (15). Familial and ethnic differences exist in the occurrence and the degree of hirsutism. The androgen receptors in the PSU are specific for the androgen dihydrotestosterone (DHT). The action of 5α-reductase is necessary to convert

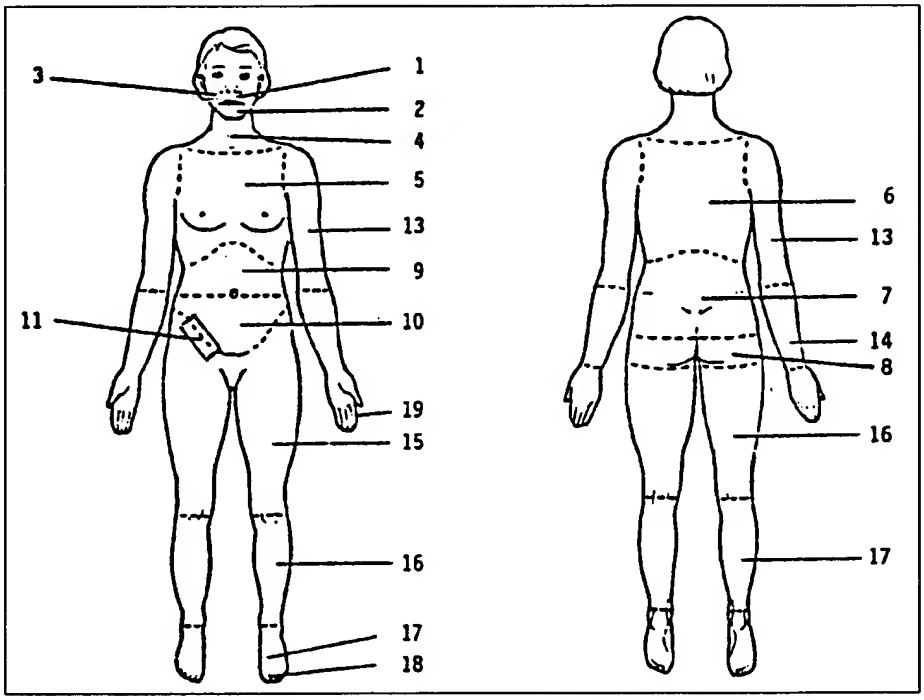


Fig. 1. Diagram showing anterior (*left*) and posterior (*right*) sites of involvement in assessment of degree and extent of hirsutism on the basis of the Ferriman-Gallwey scale. See Table 1 for definitions of hair gradings at each site. Modified from Ferriman D, Gallwey JD. *Journal of Clinical Endocrinology & Metabolism.* 1961;21:1442-1443. Table 1. ©The Endocrine Society.

Table 1
Definition of Hair Gradings at Each of 19 Sites (Ferriman-Gallwey Scale)*

Site	Grade	Definition
1 Upper lip	1	A few hairs at outer margin
	2	A small mustache at outer margin
	3	A mustache extending halfway from outer margin
	4	A mustache extending to midline
2 Chin	1	A few scattered hairs
	2	Scattered hairs with small concentrations
	3 & 4	Complete cover, light and heavy, respectively
3 Sideburns	1	Few nonterminal hairs
	2	More nonterminal hairs
	3	Terminal hair on side of face
	4	Terminal hair extending to mandible
4 Neck	1	Few hairs on neck
	2	More hairs on neck
	3 & 4	Complete cover, light and heavy, respectively
5 Chest	1	Circumareolar hairs
	2	With midline hair in addition
	3	Fusion of these areas, with three-quarter cover
	4	Complete cover
6 Upper back	1	A few scattered hairs
	2	Rather more, still scattered
	3 & 4	Complete cover, light and heavy, respectively
7 Lower back	1	A sacral tuft of hair
	2	With some lateral extension
	3 & 4	Three-quarter cover or complete cover, respectively
8 Buttocks	1 & 2	Few or many hairs, respectively, over lower buttocks
	3 & 4	Hair extending to upper buttocks, light and heavy, respectively
9 Upper abdomen	1	A few midline hairs
	2	Rather more, still midline involvement
	3 & 4	Half and full cover, respectively
10 Lower abdomen	1	A few midline hairs
	2	A midline streak of hair
	3	A midline band of hair
	4	An inverted V-shaped growth
11 Inguinal area	1	Pubic hair extending to inguinal area
	2	A few hairs below inguinal area
	3 & 4	Complete cover below inguinal area, light and heavy, respectively
12 Perianal area	1	Hair encircling introitus and anus
	2	Hair extending to inner thigh
	3 & 4	Hair on inner thigh and buttocks, light and heavy, respectively
13 Arm	1	Sparse growth affecting no more than a quarter of the limb surface
	2	More than this; cover still incomplete
	3 & 4	Complete cover, light and heavy, respectively
14 Forearm	1, 2, 3, 4	As for arm
15 Thigh	1, 2, 3, 4	As for arm
16 Leg	1, 2, 3, 4	As for arm
17 Foot	1	A few hairs on dorsum of foot
	2	More hair on dorsum of foot
	3 & 4	Hair over one-half or three-quarters or more, respectively, of dorsum
18 Toes	1 & 2	Few hairs or many hairs, respectively, on big toe
	3 & 4	Few hairs or many hairs, respectively, on other toes
19 Fingers	1	Few hairs on proximal phalanx—dorsal surface
	2	Many hairs on proximal phalanx—dorsal surface
	3	Few hairs on 2nd phalanx—dorsal surface
	4	Many hairs on 2nd phalanx—dorsal surface
Total score	_____	

*See Figure 1 for depiction of sites. Grade 0 at all sites indicates absence of terminal hair.
Modified from Ferriman and Gallwey (12).

testosterone to DHT in the hair follicle. Women with hirsutism may have increased levels of 5α-reductase. Measurement of plasma DHT, however, is not useful in assessing hirsutism, acne, or androgenetic alopecia (16).

The rapid development of hirsutism may be associated with severe hyperandrogenemia, other PSU changes, and ovarian dysfunction. Rapid onset and very high androgen levels should alert the clinician to the possibility of a neoplasm (adrenal or ovarian).

Androgenetic Alopecia

Alopecia or hair loss may be a clinical result of androgen excess. Fifteen percent of reproductive-age women with the manifestation of alopecia and no other signs of hyperandrogenism have hyperandrogenemia (17). Typically, the hair loss occurs at the vertex, but it may affect the crown later and eventually produce a diffuse pattern of hair loss. With the pronounced hyperandrogenemia found in virilizing states, one may see the typical male pattern balding including bitemporal hair loss. The latter finding is often associated with severe cutaneous manifestations of androgen excess, clitorimegaly, and menstrual dysfunction including amenorrhea.

Other factors such as genetic predisposition, nutritional deficiencies, recent weight loss, anemia, and thyroid dysfunction are responsible for the alopecia in many women without hyperandrogenemia. Certain drugs, including danazol, anabolic agents, and isotretinoin, may also cause hair loss.

Virilization

Virilization, characterized by clitoral hypertrophy, deepening of the voice, androgenic muscle development, breast atrophy, severe hirsutism, male pattern baldness, and masculine habitus, is associated with severe hyperandrogenemia attributable to adrenal or ovarian tumors, hyperthecosis, or CAH. A rapidly progressing syndrome of androgen excess, usually with pronounced oligomenorrhea or amenorrhea, should prompt the clinician to suspect a virilizing state (18).

Ovulatory Dysfunction

Women with hyperandrogenemia have various degrees of ovulatory dysfunction (19), which may lead to infertility (20,21). The ovulatory abnormalities are brought to the physician's attention and expressed clinically as menstrual dysfunction—for example, oligomenorrhea, amenorrhea, menorrhagia, metrorrhagia, pelvic pain, premenstrual dysphoric syndrome, and disturbances in fertility. The manifestations of menstrual dysfunction range from sporadic episodes of oligo-ovulation or anovulation to amenorrhea. Typically, women with ovulatory dysfunction associated with hyperandrogenemia can have a normal or delayed menarche (22), which is frequently followed by irregular menses and episodes of amenorrhea.

In addition to hyperandrogenemia, other factors may have an influence on ovulatory activity. Examples include obesity, eating disorders (manifested by various forms of anorexia or bulimia), hyperprolactinemia, hypothalamic dysfunction, and psychopharmacologic therapy. All these conditions may be associated with severe menstrual dysfunction, including amenorrhea, and must be excluded. At initial assessment, a considerable proportion of women with hyperandrogenism may have regular menses (at 4-week intervals) (6). Some, however, may have poor ovulatory activity, prolonged follicular phase, shortened luteal phase, or anovulation (20). Indeed, spontaneous ovulation and pregnancies may also occur (23). Nevertheless, fecundity, the average time for pregnancy to occur, may be delayed, and the incidence of spontaneous miscarriages is increased (24).

Metabolic and Cardiovascular Consequences of Hyperandrogenism

The relationship between the most common condition associated with hyperandrogenism—PCOS—and long-term metabolic disorders such as insulin resistance, type 2 diabetes mellitus, and dyslipidemias is now well recognized (25,26). Hypertension may also be associated with hyperandrogenemia (27). Recent data clearly demonstrate the role of hyperinsulinemia and insulin resistance in PCOS (28); these factors have been implicated in the development of atherosclerosis and a predisposition to coronary artery disease and type 2 diabetes mellitus (29). Impaired glucose tolerance and type 2 diabetes mellitus are found in 40 to 45% of patients with PCOS at the time of initial examination (30,31).

Frequently, the clinical and biochemical findings of hyperandrogenism in women are forerunners of metabolic disturbances that have a major effect on their health (32). The association with obesity also potentiates the probability of their occurrence. Some studies have also found the presence of hyperinsulinemia and insulin resistance in lean women with PCOS (33), particularly those with oligomenorrhea.

The finding of increased low-density lipoprotein cholesterol concentrations in most obese and some lean women with PCOS (30) provides a metabolic basis for the report of increased atheroma in hirsute compared with nonhirsute women undergoing angiography (34).

The focus is shifting to the possibility of thwarting the metabolic complications by early diagnosis and treatment of PCOS. Therapeutic strategies may include insulin-sensitizing agents and other modalities together with weight reduction in obese women with PCOS, which may decrease the risks of the dysmetabolic syndrome, also called syndrome X and the insulin resistance syndrome (35). The features of the dysmetabolic syndrome are listed in Table 2. To date, no large published series have defined the natural history of PCOS, but it is prudent that, with the data available, a vigorous attempt be made to reduce the risk factors for possible development of dysmetabolic syndrome.

Psychologic Dysfunction

Cystic acne, hirsutism, and alopecia, the clinical expressions of hyperandrogenemia, can have a devastating psychosocial effect in young girls and women of repro-

Table 2
Components of the Dysmetabolic Syndrome*

Dyslipidemia
Triglycerides >140 mg/dL or high-density lipoprotein cholesterol <40 mg/dL *or*
Low-density lipoprotein cholesterol particle size <260 angstroms (26 nm)
Insulin resistance
Fasting plasma glucose >110 mg/dL *or*
Type 2 diabetes mellitus
Obesity
Body mass index >25 kg/m^2 *or*
Waist-to-hip ratio >0.85 *or*
Waist circumference >100 cm
Hypertension
Blood pressure 140/90 mm Hg
Other abnormalities
High level of serum uric acid
High level of plasminogen activator inhibitor-1

*A diagnosis of dysmetabolic syndrome requires the presence of criteria for at least 2 of the first 3 components.

ductive age. These manifestations may be associated with severe anxiety and depression. In addition, symptoms resembling those frequently classified as part of the premenstrual syndrome, such as dysphoria, are frequent complaints of patients with ovulatory dysfunction, particularly when associated with hyperandrogenemia (36).

The occurrence of obesity in conjunction with hyperandrogenism can have a further negative effect on self-esteem and self-image. The fear of social rejection can make some women reclusive and may retard their development of social skills and confidence. Such women are usually helped considerably once an accurate diagnosis of the underlying endocrine and ovarian disorders is obtained. Correction of the underlying pathophysiologic condition can help ameliorate the psychologic dysfunction.

DIAGNOSTIC EVALUATION OF HYPERANDROGENISM

In most cases, the symptoms usually associated with hyperandrogenism (hirsutism, menstrual dysfunction, and infertility) are not brought to the attention of a physician for evaluation until the patient is in her late teens, early 20s, or later. Clearly, the most common causes of hyperandrogenism begin in early adolescence; therefore, every attempt should be made to diagnose and treat these conditions as early as possible. The diagnostic responsibility lies with the primary-care physicians, pediatricians, and gynecologists, particularly those who deal with "adolescent gynecology," who first come in contact with these young patients (37,38).

The diagnostic evaluation is designed to provide information about the specific androgens involved (for example, testosterone, free testosterone, and dehydroepiandrosterone sulfate [DHEAS]), the degree of hypersecretion, the organs of origin of the androgen excess (such as the ovaries or adrenal glands), and the pathogenesis of the excess androgen production. Furthermore, the evaluation must determine whether the excessive production of the androgens is due to organ dysfunction, hyperplasia, or a neoplasm. In addition, the degree of the effect of the hyperandrogenism on the integument, reproductive system, cardiovascular system, and metabolic function should be determined.

History and Physical Examination

A thorough history and physical examination provide the most important initial diagnostic information, whereas laboratory tests should usually serve to confirm the presence of hyperandrogenemia. The history should include information about the patient's age at thelarche, adrenarche, and menarche. If obesity is present, the time of onset and the progression should be noted. The character of the menstrual cycles (their frequency, duration, and occurrence of dysmenorrhea) along with the reproductive history, including miscarriages, should be evaluated. The patient should be questioned about the age at onset and the progression of the following factors: hirsutism, acne, excessive sebum, seborrhea, and alopecia. One should also note medications used and their effects on acne and hirsutism. The family history is important in determining whether other family members have hirsutism, acne, infertility, diabetes mellitus, cardiovascular disease,

dyslipidemia, or obesity. The presence of premature balding (<35 years old) in male siblings of women with hyperandrogenism has been described (39).

A complete physical examination including pelvic examination or pelvic ultrasonography (or both) is required. In addition to the usual vital signs, the measurements of height, weight, and waist circumference and the determination of the body mass index and the waist-to-hip ratio (WHR) are essential (40). The WHR (normal women, <0.8) and body mass index are important in assessing the degree of obesity in women with hyperandrogenism. A longitudinal 4-year follow-up of 32,898 women who were 55 to 69 years of age and had an increased WHR revealed a significantly increased risk of mortality attributable to coronary artery disease (41).

Particular attention should be paid to the degree and distribution of cutaneous manifestations of androgen excess (hirsutism, acne, and alopecia). The degree of hirsutism can best be documented and graded by using a system such as the one presented in Figure 1 and Table 1. In addition, the presence of clitoral hypertrophy and acanthosis nigricans (an epiphenomenon of brown and sometimes verrucous hyperpigmentation on the sides and back of the neck, axillae, submammary region, subpanniculus areas, perineum, or vulva) should be noted. Acanthosis nigricans as well as the presence of skin tags, usually in the neck area, may be indicators of insulin resistance. Examination of the thyroid gland and breasts (presence of galactorrhea) should be emphasized. A thorough pelvic examination should include an inspection of the vulva for the presence of clitoral hypertrophy and of the cervix for the state of dilatation of the cervical os and presence and type of cervical mucus. This last factor may be of considerable help in determining the patient's ovulatory status. The uterus should be examined for size and the presence of tumors. Similarly, the adnexae should be examined for the presence of masses. In the obese patient, pelvic examination can be imprecise and limited. Under these circumstances, pelvic ultrasonography is imperative for identifying pathologic changes in the pelvic organs.

Although a pelvic ultrasound study demonstrates typical subcortical ovarian cysts ranging from 5 to 10 mm and having an increased stroma in most women with PCOS, these morphologic findings are nonspecific and may be found in other diseases or in normal women (42). Similar ultrasound patterns can be found in nonclassic adult-onset CAH, hyperprolactinemia, and thyroid dysfunction and in perimenarchal girls. Thus, the ultrasound picture of a polycystic ovary should not be used as a criterion for the diagnosis of PCOS. Despite the nonspecificity of pelvic ultrasonography, it can reveal ovarian size, the nature of ovarian follicles and stroma, the state of the endometrium, the response to therapy, and the diagnosis of ovarian neoplasms (primarily dermoids) (43).

Laboratory Studies

Although the evaluation of androgen levels and their site of secretion in patients with suspected hyperandrogenism is essential, other endocrine studies, particularly as suggested by the history and physical examination, are in order. The determination of luteinizing hormone (LH), follicle-stimulating hormone (FSH), prolactin, glucose, and lipid levels is of particular interest during the initial laboratory evaluation.

The laboratory evaluation of the patient with clinical signs of hyperandrogenism necessitates specific knowledge and assessment of the endocrine testing being performed. Special care must be exercised to select a laboratory known to the physician to be proficient in androgen determinations—that is, a laboratory that has made an effort to obtain appropriately timed blood samples and has reliably determined the "normal" or the "reference" ranges for the hormones tested. Many commercial laboratories do not provide accurate, sensitive, and reproducible hormone determinations, particularly for androgens. The most challenging problems are encountered with the determination of testosterone levels in normal female patients who have testosterone levels in the low range of assay detection. Many commercial laboratories have inaccurate normal ranges for testosterone levels, as evidenced by the fact that most research publications report much lower normal-range levels for women without hyperandrogenism than do commercial laboratories. Studies have shown that the same testosterone value measured by different laboratories may be designated as normal to abnormally high when compared with the supplied reference ranges (44,45). Thus, an acceptable clinical laboratory should provide precise results, a reliable reference range, and quality control that ensures stable values over extended periods. These features are crucial because long-term therapy for hyperandrogenism is frequently necessary.

During the initial assessment of a patient with suspected hyperandrogenism, the following plasma hormone concentrations could be determined during fasting in a specimen obtained during the first 7 days of the menstrual cycle: total and free testosterone, DHEAS, prolactin, LH and FSH, and 17-hydroxyprogesterone (17-OHP).

The determination of plasma free testosterone may help detect the presence of subtle hyperandrogenemia in some patients in whom the total testosterone level is normal. Of note, obese women may have a high free testosterone value because of reduced sex hormone-binding globulin (SHBG) attributable to hyperinsulinemia (46). A high DHEAS level may indicate an adrenal factor in androgen production and, if substantially elevated, the presence of an adrenocortical neoplasm. The usefulness of a high LH level or an elevated LH/FSH ratio (>2.5) in diagnosing the presence of polycystic ovarian state in women with hyperandrogenism is limited, inasmuch as at least a third of the patients with PCOS do not demonstrate this abnormality (47). In reference to reproductive function, however, a high LH level is associated with a poor response to attempted induction of ovulation and an increased risk of miscarriage (48).

Measurement of plasma 17-OHP levels is useful in diagnosing adrenal hyperandrogenemia as a result of 21-hydroxylase deficiency. A high 17-OHP level suggests the presence of this enzyme deficiency disease. With a suspi-

cious family history of CAH or in Ashkenazi women, the use of a cosyntropin stimulation test may be necessary to exclude this enzyme deficiency (49).

Many clinicians have found that women with clinical signs of hyperandrogenism who undergo initial assessment after the age of 35 years may not demonstrate hyperandrogenemia (Futterweit W. Unpublished data). This finding may prevail because, with aging, the hyperandrogenic activity (particularly that of the ovaries) decreases and the hirsutism will persist. Therefore, follow-up monitoring of such women for the development of associated metabolic and cardiovascular sequelae is important.

DETERMINATION OF SITE OF ANDROGEN PRODUCTION

The hypersecretion of androgens can be attributed to the ovaries, the adrenal glands, or the peripheral conversion of androgen precursors. From the standpoint of therapeutic strategy, determining the degree of contribution of the androgens originating from each site is critical. For this purpose, several endocrine function tests are available.

Adrenal Suppression Tests

Two-Day Suppression Test

The following steps are involved in the 2-day dexamethasone suppression test (50).

Day 1

- Beginning at 8 to 9 AM, obtain three blood samples at 20-minute intervals. Equal aliquots of each sample are pooled to provide the baseline sample (before administration of dexamethasone).
- Dispense to the patient eight 0.5-mg dexamethasone tablets.
- The patient takes one tablet of dexamethasone at lunch, dinner, and bedtime.

Day 2

- The patient takes one tablet of dexamethasone at breakfast, lunch, dinner, and bedtime.

Day 3

- The patient takes one tablet of dexamethasone at breakfast.
- Subsequently, beginning within 2 hours, three blood samples are obtained every 20 minutes. Equal aliquots of each sample are pooled to provide the "postdexamethasone" sample.
- Testosterone, DHEAS, and cortisol are measured on pooled blood samples obtained before and after administration of dexamethasone.

Alternative Suppression Tests

Some authorities suggest the use of 4 to 6 days of suppression with dexamethasone (51,52). These tests are conducted in the same fashion as the 2-day suppression test. Some investigators believe that such tests yield no substantially greater amount of information than the 2-day dexamethasone suppression study.

Interpretations

If the high level of testosterone is suppressed more than 40% and DHEAS is suppressed more than 60% after administration of dexamethasone, the source of the increased androgen is most likely the adrenal glands (19).

If the testosterone level fails to be suppressed but DHEAS and cortisol respond, the source of testosterone is primarily the ovaries.

If testosterone suppression is less than 40%, a combined ovarian and adrenal contribution is responsible for the excessive testosterone secretion.

If lack of androgen suppression is associated with a lack of cortisol suppression, the patient most likely has adrenal hyperfunction (for example, Cushing's disorder or adrenal cancer) or the patient may have failed to take the dexamethasone tablets as directed. Appropriate steps must be taken to determine which of these two possibilities is the likely explanation.

Stimulation Tests

Adrenal Stimulation

The cosyntropin stimulation test is conducted for detection of the various steroidogenic enzyme deficiencies of the adrenal glands, predominantly 21-hydroxylase deficiency (49). This test is indicated only if a screening morning 17-OHP level is above normal or the suspicion of such an enzymatic deficiency is high, as in Ashkenazi and Hispanic women. A morning plasma cortisol level and 17-OHP level are determined, and 0.25 mg of cosyntropin is administered intravenously. Blood sampling is repeated at 60 minutes. Specific testing for other adrenocortical defects may also be performed to exclude the presence of 3β-ol dehydrogenase and 11β-hydroxylase deficiencies (49).

Ovarian Stimulation

The gonadotropin-releasing hormone (GnRH) stimulation test (53) is conducted to confirm the ovarian origin of the excessive androgens. This test involves suppression of the adrenal glands with dexamethasone and, while the adrenal glands are suppressed, stimulation of the ovaries with a GnRH agonist (for example, nafarelin). Blood samples are collected at various time intervals during a 25-hour period. The hypersecretion of 17-OHP is the endpoint of the test, indicating ovarian involvement. This test is useful only in clinical research of PCOS to identify the presence of 17-hydroxylase dysfunction.

DETERMINATION OF DEGREE OF ANDROGEN-INDUCED CHANGES

Integument

The degree and distribution of acne and hirsutism, as well as alopecia when present, should be assessed with use of the Ferriman-Gallwey scoring technique (Fig. 1 and Table 1) or a similar scale (12-14).

Ovaries

Menstrual cycles can be defined as ovulatory or anovulatory. A simple method to evaluate ovulation is the use of a basal body temperature chart, in which a biphasic pattern suggests an ovulatory cycle, whereas a flat pattern is consistent with anovulation. A serum progesterone level of less than 2 ng/mL after the 21st day of the menstrual cycle or after a perceived increase in the basal body temperature is also consistent with ovulatory dysfunction.

Metabolic Sequelae

The metabolic sequelae of hyperandrogenemia are of major importance and need to be considered during the early assessment of patients with clinical signs of hyperandrogenism. The increased incidences of hyperinsulinism and insulin resistance make these patients susceptible to impaired glucose tolerance and, ultimately, overt type 2 diabetes mellitus. Although the "gold standard" for defining insulin resistance (other than a euglycemic-hyperinsulinemic clamp study) is the frequently sampled intravenous glucose tolerance test, fasting morning glucose and insulin measurements may be a useful screening test to determine the presence of these metabolic abnormalities. A ratio of fasting glucose/insulin <4.5 has been proposed as an indicator of insulin resistance in such patients (54). In a recently published study, reevaluation of this issue provided evidence that an effective method to determine the presence of insulin resistance is to measure the total integrated insulin response to orally administered glucose (55). The fasting insulin level correlated highly with this more elaborate testing.

In addition to impaired glucose metabolism, dyslipidemia characterized by reduced high-density lipoprotein cholesterol and increased triglyceride levels is frequently noted (56). These manifestations place patients with hyperandrogenism at an increased risk for later cardiovascular atherogenic complications. Although treatment of hyperandrogenism is important for the alleviation of the cutaneous (acne and hirsutism) or menstrual (ovulatory dysfunction) abnormalities, a major goal of treatment is possible amelioration of the metabolic disorders. The natural history of the disease has not been studied in a sufficiently large series of patients with hyperandrogenism. On the basis of the available data, however, these patients are clearly at risk, and this factor should be at the forefront of therapeutic considerations.

THERAPY FOR HYPERANDROGENISM

Although the therapeutic approach for hyperandrogenism is usually directed at the primary symptom of the patient—acne, hirsutism, alopecia, menstrual dysfunction, infertility, or an associated metabolic disorder—the primary concern is a thorough diagnostic evaluation to determine the degree of hyperandrogenism, the site of excess androgen production, and the presence of all pathophysiologic and metabolic manifestations. Early recognition of the disease process and timely therapeutic intervention should be of foremost concern. Frequently, the disease can be diagnosed in the perimenarchal or early postmenarchal period of the female patient's development. Early treatment may prevent serious problems with acne, hirsutism, dysfunctional bleeding, and infertility and possibly ameliorate the potential later development of metabolic and cardiovascular complications (37).

Therapy may be categorized on the basis of the results of dynamic tests and the pathophysiologic findings:

- Predominantly glucocorticoid-suppressible hyperandrogenism (probably adrenocortical in origin) (57)
- Combined ovarian and adrenal hyperandrogenism
- Predominantly ovarian hyperandrogenism
- Rare forms of hyperandrogenism:
  Congenital adrenal hyperplasia
  Adrenal tumors
  Ovarian tumors
  Hyperthecosis

The pharmacologic options in the treatment of the aforementioned disorders include the following:

- Suppression of adrenal androgens (administration of glucocorticoids, usually in physiologic doses of dexamethasone or prednisone)
- Suppression of ovarian androgens (administration of female sex steroids, in the form of either birth control pills or estrogens and progestins, or administration of gonadotropin secretion blocking agents—that is, GnRH agonists with "add-back" estrogen therapy)
- Treatment with antiandrogens (for example, spironolactone and flutamide)
- Treatment with insulin-sensitizing agents (metformin and thiazolidinediones)
- Treatment with 5α-reductase-inhibiting agents (finasteride)
- Bromocriptine

Of importance, treatment with antiandrogens, insulin-sensitizing agents, and 5α-reductase-inhibiting agents is not specifically approved by the US Food and Drug Administration and is contraindicated during pregnancy.

Nonpharmacologic interventions for hyperandrogenic states include the following: (*1*) weight reduction, (*2*)

surgical excision of a virilizing adrenal or ovarian tumor, and (*3*) electrocautery in patients with PCOS.

Treatment With Glucocorticoids

Adrenal hyperandrogenism responds well to glucocorticoid therapy with prednisone or dexamethasone (58-60). Usually, 5 to 7.5 mg of prednisone or 0.25 to 0.5 mg of dexamethasone is administered daily after supper for 2 or 3 months. This treatment frequently results in normalization of the androgens. If the androgen levels are normalized, the dose is lowered to 5 mg of prednisone or 0.25 mg of dexamethasone for 2 to 3 months, at which time the dose can be halved or totally discontinued. The androgen levels should then be assessed every 3 to 4 months for 1 year for recurrence of hyperandrogenemia. Of note, the possibility of recurrence of high testosterone levels is greater than recurrence of high DHEAS levels. In most instances, DHEAS levels remain suppressed indefinitely after treatment (60). Suppression of androgens usually results in sustained amelioration of acne and has a minor effect on hirsutism (slowing of growth rate and softening of hair). Improvement in ovulatory activity, return of fertility, and increased sensitivity to the effect of clomiphene citrate on induction of ovulation often occur with glucocorticoid therapy (17,58,60).

The long-term adverse effects of minimal-dose glucocorticoid therapy on bone resorption and dysmetabolic syndrome have not been scientifically studied. Routine follow-up studies of patients have not shown significant effects on glucose or lipid metabolism; however, short-acting corticosteroids such as prednisone would be less likely to pose such risks in comparison with the use of dexamethasone.

Treatment With Oral Contraceptives

Oral contraceptives (OCs) are used widely for treatment of hyperandrogenism (61-63). The use of third-generation OCs in the treatment of ovarian hyperandrogenism appears promising in that little if any androgenic effect has been noted with desogestrel and norgestimate, the progestins used in these OC agents (64). Acne and mild hirsutism are often decreased. The combination of OCs and antiandrogenic agents, particularly spironolactone, is often used in patients with moderate to severe hirsutism or alopecia (65). A major benefit is the reduction in the incidence of endometrial and ovarian cancer in patients who use OCs (66). Contraindications to their use may be a history of phlebitis, severe migraine, substantial weight gain, and the risk of increased insulin resistance. Long-term use may mask severe ovulatory dysfunction, which may progress to anovulation and amenorrheic states that are more resistant to induction of ovulation.

Treatment With Antiandrogens

Spironolactone

Spironolactone is an antiandrogen that competes with testosterone and DHT at the androgen receptor level. The minimal dose should be 100 mg daily in divided dosage and may be increased to 200 mg daily as tolerated. The combination of spironolactone and OCs is frequently used (67). The LH suppressive effect of OCs makes this combination treatment more effective than spironolactone monotherapy. This combined drug strategy minimizes the frequently noted polymenorrhea when spironolactone is used alone. Some of the side effects include light-headedness, fatigue, mood swings, reduced libido, headaches, and mastalgia. In patients with androgenetic alopecia, the use of spironolactone is effective in improving the rate of hair regrowth and preventing further scalp hair loss.

Cyproterone Acetate

Cyproterone acetate has not been approved by the US Food and Drug Administration. Reports indicate that this progestational antiandrogen is effective in the treatment of hyperandrogenism. When administered in conjunction with ethinyl estradiol, it is equal or perhaps slightly superior to combination therapy with spironolactone and OCs (68).

Flutamide

Flutamide, an antiandrogen that blocks androgen uptake and nuclear binding, is a very effective drug in treating hyperandrogenism, but it has the potential, albeit infrequent, adverse effect of fatal hepatotoxicity. Flutamide should be used cautiously only in the most severe cases resistant to other forms of treatment (69). Recent data indicate that a dosage of 250 mg daily may be as effective as 250 mg twice a day.

Cimetidine

An H_2 blocker, cimetidine has not been found to be useful in the treatment of hyperandrogenism and is not recommended in this setting.

Ketoconazole

Ketoconazole is an antifungal agent that has shown some effectiveness in treating symptoms of androgen excess. Reports of serious hepatotoxicity, however, have made this drug treatment of questionable value.

Treatment With 5α-Reductase Inhibitors

Finasteride is a 5α-reductase inhibitor that blocks the intracellular conversion of testosterone to DHT; thus, the amount of DHT available for interacting with the androgen receptor is reduced. Finasteride has a predominant effect on the type 2 isoenzyme of 5α-reductase that specifically affects sebaceous gland activity. Reports of its use in women have been limited to those with hirsutism, and it seems to be comparable to spironolactone in reducing anagen hair shaft diameter when administered in a daily dose of 5 to 7.5 mg. It is associated with minimal gastrointestinal side effects, and it does not alter menstrual cyclicity. The plasma levels of testosterone may increase during treatment, whereas the DHT level decreases (70,71). Of utmost importance, the patient should be aware that she

must avoid pregnancy during treatment with finasteride because of the potential for causing ambiguous genitalia in a male fetus.

Treatment With Insulin-Sensitizing Agents

The advent of drugs that enhance insulin sensitivity and action makes agents such as metformin and the thiazolidinediones important in the therapeutic strategy for patients with hyperandrogenism and some of the associated metabolic disturbances—specifically, insulin resistance (72,73). Studies have indicated that reducing plasma insulin levels substantially ameliorates the hyperandrogenism of PCOS. Insulin-sensitizing agents may become a choice for initial treatment of women with hyperandrogenism, particularly those with PCOS and insulin resistance, obesity, and moderate to severe oligomenorrhea.

Metformin

Reduction in the manifestations of hyperandrogenism and improved menstrual function have been reported with administration of metformin. Metformin has been administered in conjunction with ovulation-inducing agents, such as clomiphene citrate, and has not been reported to be teratogenic (74).

Many studies have demonstrated the most dramatic reduction in hyperandrogenism in obese subjects with PCOS; hence, the issue has been raised whether weight reduction alone accounts for this effect. Questions have been posed about the reported efficacy of metformin in increasing insulin sensitivity in this type of patient because of the small numbers of subjects and questionable study designs in most reports dealing with this issue (75). Recently, however, a carefully designed study suggested a definite positive effect of metformin in a relatively large population of patients (76).

The recommended dosage of metformin for treatment of hyperandrogenic states is the initiation of therapy with 850 mg (one tablet) in the morning with breakfast, and then increasing the dosage to 1,700 mg after 2 to 3 weeks in divided doses with breakfast and dinner. Alternatively, metformin therapy can be initiated at 500 mg with dinner and increased to 500 mg twice and three times daily or to 1,000 mg twice a day as tolerated. The most common side effects are gastrointestinal; they consist of bloating, nausea, vomiting, and diarrhea and frequently occur during initiation of treatment. The occurrence of lactic acidosis is extremely rare and is more likely in those patients with renal impairment. Use of metformin should be discontinued before administration of contrast dye, particularly in subjects with decreased renal function (77).

Thiazolidinediones

Thiazolidinediones have been used alone or in combination in the treatment of type 2 diabetes mellitus. The primary effect of these drugs is achieved by improving insulin sensitivity in muscle and adipose tissue as well as inhibiting hepatic gluconeogenesis. The decrease in insulin resistance is accompanied by a reduction in hyperinsulinemia, without associated weight changes. Studies have indicated that the use of troglitazone improved insulin sensitivity and subsequently decreased insulin-mediated ovarian androgen excess in women with PCOS, even in those subjects with severe obesity (78,79). In view of the occasional reports of serious hepatic disturbances associated with its use, however, troglitazone was removed from the market. Other agents in the thiazolidinedione class of drugs may prove to be effective in the treatment of PCOS, but to date, no scientific studies have been published of their use in PCOS.

Comment

Overall, the use of insulin-sensitizing agents has demonstrated clinical improvement in hirsutism and menstrual cyclicity. Improvement in responsiveness to ovulation-inducing agents (specifically, clomiphene citrate) has been reported (74). The effects on improvement of the cutaneous manifestations of hyperandrogenism may not be as pronounced as those achieved with antiandrogen treatment in combination with OCs.

Further studies are needed to determine the exact role of metformin and thiazolidinediones as therapeutic agents, or possibly initial monotherapy, in the treatment of hyperandrogenic states, including PCOS.

Treatment With GnRH Agonists

The use of GnRH agonists is most effective in severe forms of ovarian hyperandrogenism. Depot preparations of GnRH agonists may be administered at monthly intervals. Because of the severe hypoestrogenemia induced by GnRH agonists, however, a concurrent add-back therapy with estrogen and progesterone (the latter when the uterus is present) is essential (80,81). This therapy will also correct the severe vasomotor symptoms and other side effects of the resulting hypoestrogenemia. GnRH agonists have shown no effect in reducing hyperinsulinism in instances of ovarian hyperandrogenism.

Treatment With Bromocriptine

The use of bromocriptine, a dopamine receptor agonist, in a divided dosage of 5 to 7.5 mg daily with meals is indicated in the subset of women with hyperandrogenism who have hyperprolactinemia. No convincing data indicate that this treatment is effective in women with hyperandrogenism who do not have high levels of circulating prolactin (78,79). Bromocriptine improves menstrual cyclicity in patients with hyperprolactinemia who have PCOS (82) and may reduce some of the associated hirsutism, which is related to augmented production of adrenal androgen attributable to the hyperprolactinemia. Treatment with bromocriptine should be initiated gradually so as to minimize initial light-headedness, hypotension, and nausea (vaginal or rectal administration can also reduce symptoms). Alternatively, use of cabergoline in dosages of 0.5 mg weekly or twice weekly may be associated with fewer side effects.

Weight Reduction

Obesity is present in 55 to 65% of patients with PCOS, the most common form of hyperandrogenism, and is associated with an increase in the WHR and a history of perimenarchal onset. Obesity is frequently associated with insulin resistance and hyperinsulinism, which act synergistically with LH to intensify ovarian hyperandrogenism, decrease hepatic production of SHBG, and reduce insulin-like growth factor binding protein-1 (83). Moreover, obesity reinforces the genetic predisposition to hormonal and ovulatory disorders in PCOS. Metabolic complications are more common in obese patients with PCOS (impaired glucose tolerance in approximately 40% of subjects with obesity, hypertension, dyslipidemias, and possible development of estrogen-dependent tumors) than in lean women with PCOS.

Weight loss in patients with hyperandrogenism, with or without the clinical presence of PCOS, should be the first therapeutic option because it decreases androgen levels, increases SHBG, and may restore ovulation. As little as a 7% reduction in body weight can restore fertility, decrease hirsutism in some women with androgen excess, and improve the response to induction of ovulation (84,85). Investigators have also demonstrated that abnormalities in 17,20-lyase activity diminish in parallel with reduction of hyperinsulinemia (86).

Individualization of Therapy

The choice of therapeutic intervention for hyperandrogenism depends on several factors:

- The source of the hyperandrogenism
- The goal of therapy
- The long-term risks and benefits

The evaluation and treatment of hyperandrogenism and hyperandrogenemia should begin as early as possible after the onset of symptoms. Often, mild symptoms, such as menstrual irregularities, hirsutism, and acne, are overlooked or discounted, and the result is their progression. For example, adolescent patients diagnosed with a hyperandrogenic disorder should undergo assessment for treatment based on the cause of the hyperandrogenism, the long-term goals, and the potential benefits and risks.

For those patients with glucocorticoid-suppressible hyperandrogenism, successful treatment for 1 to 2 years may yield a long-term remission and prevent metabolic and ovulatory disorders (60).

In non-glucocorticoid-suppressible hyperandrogenism, the use of an OC-antiandrogen combination may be effective in treating the hyperandrogenemia but may not ameliorate the long-term metabolic and ovulatory abnormalities. Therefore, patients should be counseled that discontinuation of therapy may lead to recurrence of the hyperandrogenic problems.

Insulin-sensitizing drugs have been shown to be effective in short-term studies, but long-term outcomes are as yet unknown. Those patients, regardless of age, with severe forms of PCOS, such as those with insulin resistance, obesity, and acanthosis nigricans, are candidates for this therapeutic modality.

CONCLUSION

The primary responsibility of the physician in the assessment of a patient with symptoms associated with hyperandrogenism—for example, acne, hirsutism, reproductive disorders, or metabolic diseases—must be to determine the presence of high levels of androgens and their source. Careful attention should be focused on whether damage to organ systems other than the one responsible for the patient's major complaint is also present. Early diagnosis is critical to an effective therapeutic strategy. The adolescent patient with hyperandrogenism may not undergo thorough assessment before being treated symptomatically for irregular menses, hirsutism, or acne. Many young women are given oral contraceptives to "regulate menstrual cycles" or to "improve acne" without a careful determination of the source of the hyperandrogenism and selection of the most appropriate therapy for maximizing the long-term benefits (37,38).

Adequate diagnostic evaluation hinges on the availability of a laboratory capable of accurate measurement of low levels of androgens, particularly testosterone, and the ability of the laboratory to provide reproducible results. The physician should ensure that the laboratory performing these highly specialized hormone assays provides accurate and reproducible results.

The initial goal of therapy is to lower the androgen levels, which should occur within the first 2 or 3 months of therapy. Once this goal has been achieved, evaluation of the degree of the return of function in organs originally impaired by hyperandrogenism should be undertaken. Then, if necessary, direct therapy toward further improvement of these functions can be instituted.

Management of non-tumor-related hyperandrogenemia and polycystic ovaries necessitates careful long-range monitoring by an endocrinologist for proper maintenance of a euandrogenemic state. Vigilance must be exercised to detect any signs of development of the metabolic consequences of hyperandrogenism, including diabetes mellitus, hypertension, dyslipidemias, or atherosclerosis.

REFERENCES

1. **Haseltine F, Redmond GP, Colston A, Wild R.** Androgens and women's health. *Clinician (NICHD).* August 1994;12(No. 3):1-32.
2. **Held BL, Nader S, Rodriguez-Rigau LJ, Smith KD, Steinberger E.** Acne and hyperandrogenism. *J Am Acad Dermatol.* 1984;10(2 Pt 1):223-226.
3. **Lucky AW.** Hormonal correlates of acne and hirsutism. *Am J Med.* 1995;98(1A):89S-94S.
4. **Rosenfield RL, Deplewski D.** Role of androgens in the developmental biology of the pilosebaceous unit. *Am J Med.* 1995;98(1A):80S-88S.
5. **Slayden SM, Azziz R.** The role of androgen excess in acne. In: Azziz R, Nestler JE, Dewailly D, eds. *Androgen Excess Disorders in Women.* Philadelphia: Lippincott-Raven, 1997: 131-140.

6. **Steinberger E, Rodriguez-Rigau LJ, Smith KD, Held B.** The menstrual cycle and plasma testosterone levels in women with acne. *J Am Acad Dermatol.* 1981;4:54-58.
7. **Peserico A, Angeloni G, Bertoli P, et al.** Prevalence of polycystic ovaries in women with acne. *Arch Dermatol Res.* 1989;281:502-503.
8. **Ravin JG, Hodge GP.** Hypertrichosis portrayed in art. *JAMA.* 1969;207:533-534.
9. **Ewing JA, Rouse BA.** Hirsutism, race and testosterone levels: comparison of East Asians and Euroamericans. *Hum Biol.* 1978;50:209-215.
10. **Balducci R, Toscano V.** Bioactive and peripheral androgens in prepubertal simple hypertrichosis. *Clin Endocrinol (Oxf).* 1990;33:407-414.
11. **Lunde O.** A study of body hair density and distribution in normal women. *Am J Phys Anthropol.* 1984;64:179-184.
12. **Ferriman D, Gallwey JD.** Clinical assessment of body hair growth in women. *J Clin Endocrinol Metab.* 1961;21: 1440-1447.
13. **Hatch R, Rosenfield RL, Kim MH, Tredway D.** Hirsutism: implications, etiology, and management. *Am J Obstet Gynecol.* 1981;140:815-830.
14. **Redmond GP.** *Androgenic Disorders.* New York: Raven Press, 1995: 611-628.
15. **Goldzieher JW, Young RL.** Selected aspects of polycystic ovarian disease. *Endocrinol Metab Clin North Am.* 1992;21:141-171.
16. **Horton R, Hawks D, Lobo R.** 3 Alpha, 17 beta-androstanediol glucuronide in plasma: a marker of androgen action in idiopathic hirsutism. *J Clin Invest.* 1982;69: 1203-1206.
17. **Futterweit W, Dunaif A, Yeh HC, Kingsley P.** The prevalence of hyperandrogenism in 109 consecutive female patients with diffuse alopecia. *J Am Acad Dermatol.* 1988;19(5 Pt 1):831-836.
18. **Futterweit W.** Clinical evaluation of androgen excess. In: Azziz R, Nestler JE, Dewailly D, eds. *Androgen Excess Disorders in Women.* Philadelphia: Lippincott-Raven, 1997: 625-631.
19. **Steinberger E, Smith KD, Rodriguez-Rigau LJ.** Testosterone, dehydroepiandrosterone, and dehydroepiandrosterone sulfate in hyperandrogenic women. *J Clin Endocrinol Metab.* 1984;59:471-477.
20. **Smith KD, Rodriguez-Rigau LJ, Tcholakian RK, Steinberger E.** The relation between plasma testosterone levels and the lengths of phases of the menstrual cycle. *Fertil Steril.* 1979;32:403-407.
21. **Steinberger E, Smith KD, Tcholakian RK, Rodriguez-Rigau LJ.** Testosterone levels in female partners of infertile couples: relationship between androgen levels in the woman, the male factor, and the incidence of pregnancy. *Am J Obstet Gynecol.* 1979;133:133-138.
22. **Futterweit W.** Polycystic ovary syndrome: clinical perspectives and management. *Obstet Gynecol Surv.* 1999;54:403-413.
23. **Goldzieher JW, Green JA.** The polycystic ovary. I. Clinical and histological findings. *J Clin Endocrinol Metab.* 1962;22:325-338.
24. **Sagle M, Bishop K, Ridley N, et al.** Recurrent early miscarriage and polycystic ovaries. *BMJ.* 1988;297:1027-1028.
25. **Wild RA.** Obesity, lipids, cardiovascular risk, and androgen excess. *Am J Med.* 1995;98(1A):27S-32S.
26. **Dunaif A.** Hyperandrogenic anovulation (PCOS): a unique disorder of insulin action associated with an increased risk of non-insulin-dependent diabetes mellitus. *Am J Med.* 1995;98(1A):33S-39S.
27. **Ayala C, Steinberger E, Sweeney A, et al.** The relationship of serum androgens and ovulatory status to blood pressure in reproductive-age women. *Am J Hypertens.* 1999;12(8 Pt 1):772-777.
28. **Dunaif A, Segal KR, Futterweit W, Dobrjansky A.** Profound peripheral insulin resistance, independent of obesity, in polycystic ovary syndrome. *Diabetes.* 1989;38: 1165-1174.
29. **Conway GS, Agrawal R, Betteridge DJ, Jacobs HS.** Risk factors for coronary artery disease in lean and obese women with the polycystic ovary syndrome. *Clin Endocrinol (Oxf).* 1992;37:119-125.
30. **Legro RS, Kunselman AR, Dodson WC, Dunaif A.** Prevalence and predictors of risk for type 2 diabetes mellitus and impaired glucose tolerance in polycystic ovary syndrome: a prospective, controlled study in 254 affected women. *J Clin Endocrinol Metab.* 1999;84:165-169.
31. **Ehrmann DA, Barnes RB, Rosenfield RL, Cavaghan MK, Imperial J.** Prevalence of impaired glucose tolerance and diabetes in women with polycystic ovary syndrome. *Diabetes Care.* 1999;22:141-146.
32. **Dahlgren E, Johansson S, Lindstedt G, et al.** Women with polycystic ovary syndrome wedge resected in 1956 to 1965: a long-term follow-up focusing on natural history and circulating hormones. *Fertil Steril.* 1992;57:505-513.
33. **Nestler JE, Jakubowicz DJ.** Lean women with polycystic ovary syndrome respond to insulin reduction with decreases in ovarian P450c17 alpha activity and serum androgens. *J Clin Endocrinol Metab.* 1997;82:4075-4079.
34. **Wild RA, Grubb B, Hartz A, Van Nort JJ, Bachman W, Bartholomew M.** Clinical signs of androgen excess as risk factors for coronary artery disease. *Fertil Steril.* 1990;54: 255-259.
35. **Fagan TC, Deedwania PC.** The cardiovascular dysmetabolic syndrome. *Am J Med.* 1998;105(1A):77S-82S.
36. **Cronin L, Guyatt G, Griffith L, et al.** Development of a health-related quality-of-life questionnaire (PCOSQ) for women with the polycystic ovary syndrome (PCOS). *J Clin Endocrinol Metab.* 1998;83:1976-1987.
37. **Smith KD, Rodriguez-Rigau LJ, Steinberger E.** Preventive reproductive medicine. In: Steinberger E, Frajese G, Steinberger A, eds. *Reproductive Medicine—Serono Symposia Publications for Raven Press.* Vol 29. New York: Raven Press, 1986: 337-344.
38. **Steinberger E, Rodriguez-Rigau LJ, Ayala C, Smith KD.** Consequences of hyperandrogenism during adolescence on the ovarian function of the adult female. In: Steinberger E, Frajese G, Steinberger A, eds. *Reproductive Medicine—Serono Symposia Publications for Raven Press.* Vol 93. New York: Raven Press, 1993: 253-264.
39. **Ferriman D, Purdie AW.** The inheritance of polycystic ovarian disease and possible relationship to premature balding. *Clin Endocrinol (Oxf).* 1979;11:291-300.
40. **AACE/ACE Obesity Task Force.** AACE/ACE position statement on the prevention, diagnosis, and treatment of obesity (1998 revision). *Endocr Pract.* 1998;4:297-350.
41. **Prineas RJ, Folsom AR, Kaye SA.** Central adiposity and increased risk of coronary artery disease mortality in older women. *Ann Epidemiol.* 1993;3:35-41.
42. **Polson DW, Adams J, Wadsworth J, Franks S.** Polycystic ovaries—a common finding in normal women. *Lancet.* 1988;1:870-872.
43. **Futterweit W, Scher J, Nunez AE, Strauss L, Rayfield EJ.** A case of bilateral dermoid cysts, insulin resistance, and polycystic ovarian disease: association of ovarian tumors with polycystic ovaries with review of the literature. *Mt Sinai J Med.* 1983;50:251-255.

44. **Steinberger E, Ayala C, Hsi B, et al.** Utilization of commercial laboratory results in management of hyperandrogenism in women. *Endocr Pract.* 1998;4:1-10.
45. **Ayala C, Steinberger E, Smith KD, Rodriguez-Rigau LJ, Petak SM.** Serum testosterone levels and reference ranges in reproductive-age women. *Endocr Pract.* 1999;5: 322-329.
46. **Penttila TL, Koskinen P, Penttila TA, Anttila L, Irjala K.** Obesity regulates bioavailable testosterone levels in women with or without polycystic ovary syndrome. *Fertil Steril.* 1999;71:457-461.
47. **Steinberger E, Smith KD, Rodriguez-Rigau LJ.** Hyperandrogenism and female infertility. In: Crogsignani PG, Rubin B, eds. *Endocrinology of Human Infertility: New Aspects.* London: Academic Press, Inc., 1981: 327-342.
48. **Shoham Z, Jacobs HS, Insler V.** Luteinizing hormone: its role, mechanism of action, and detrimental effects when hypersecreted during the follicular phase. *Fertil Steril.* 1993;59:1153-1161.
49. **New MI, Lorenzen F, Lerner AJ, et al.** Genotyping steroid 21-hydroxylase deficiency: hormonal reference data. *J Clin Endocrinol Metab.* 1983;57:320-326.
50. **Steinberger E, Rodriguez-Rigau LJ, Smith KD.** The prognostic value of acute adrenal suppression and stimulation tests in hyperandrogenic women. *Fertil Steril.* 1982;37:187-192.
51. **Abraham GE, Maroulis GB, Buster JE, Chang RJ, Marshall JR.** Effect of dexamethasone on serum cortisol and androgen levels in hirsute patients. *Obstet Gynecol.* 1976;47:395-402.
52. **Abraham GE, Chakmakjian ZH, Buster JE, Marshall JR.** Ovarian and adrenal contributions to peripheral androgens in hirsute women. *Obstet Gynecol.* 1975;46:169-173.
53. **Barnes RB, Rosenfield RL, Burstein S, Ehrmann DA.** Pituitary-ovarian responses to nafarelin testing in the polycystic ovary syndrome. *N Engl J Med.* 1989;320:559-565.
54. **Legro RS, Finegood D, Dunaif A.** A fasting glucose to insulin ratio is a useful measure of insulin sensitivity in women with polycystic ovary syndrome. *J Clin Endocrinol Metab.* 1998;83:2694-2698.
55. **Yeni-Komshian H, Carantoni M, Abbasi F, Reaven GM.** Relationship between several surrogate estimates of insulin resistance and quantification of insulin-mediated glucose disposal in 490 healthy nondiabetic volunteers. *Diabetes Care.* 2000;23:171-175.
56. **Wild RA, Applebaum-Bowden D, Demers LM, et al.** Lipoprotein lipids in women with androgen excess: independent associations with increased insulin and androgen. *Clin Chem.* 1990;36:283-289.
57. **Carmina E.** Prevalence of adrenal androgen excess in PCOS. In: Azziz R, Nestler JE, Dewailly D, eds. *Androgen Excess Disorders in Women.* Philadelphia: Lippincott-Raven, 1997: 385-393.
58. **Rodriguez-Rigau LJ, Smith KD, Tcholakian RK, Steinberger E.** Effect of prednisone on plasma testosterone levels and on duration of phases of the menstrual cycle in hyperandrogenic women. *Fertil Steril.* 1979;32: 408-413.
59. **Nader S, Rodriguez-Rigau LJ, Smith KD, Steinberger E.** Acne and hyperandrogenism: impact of lowering androgen levels with glucocorticoid treatment. *J Am Acad Dermatol.* 1984;11(2 Pt 1):256-259.
60. **Steinberger E, Rodriguez-Rigau LJ, Petak SM, Weidman ER, Smith KD, Ayala C.** Glucocorticoid therapy in hyperandrogenism. *Baillieres Clin Obstet Gynaecol.* 1990;4:457-471.
61. **Azziz R, Gay F.** The treatment of hyperandrogenism with oral contraceptives. *Semin Reprod Endocrinol.* 1989;7: 246-254.
62. **van der Vange N, Blankenstein MA, Kloosterboer HJ, Haspels AA, Thijssen JH.** Effects of seven low-dose combined oral contraceptives on sex hormone binding globulin, corticosteroid binding globulin, total and free testosterone. *Contraception.* 1990;41:345-352.
63. **London RS, Chapdelaine A, Upmalis D, Olson W, Smith J.** Comparative contraceptive efficacy and mechanism of action of the norgestimate-containing triphasic oral contraceptive. *Acta Obstet Gynecol Scand Suppl.* 1992; 156:9-14.
64. **Redmond GP, Olson WH, Lippman JS, Kafrissen ME, Jones TM, Jorizzo JL.** Norgestimate and ethinyl estradiol in the treatment of acne vulgaris: a randomized, placebo-controlled trial. *Obstet Gynecol.* 1997;89:615-622.
65. **Cumming DC, Yang JC, Rebar RW, Yen SS.** Treatment of hirsutism with spironolactone. *JAMA.* 1982;247:1295-1298.
66. **Redmond GP.** Treatment of androgenic disorders. In: Redmond GP, ed. *Androgenic Disorders.* New York: Raven Press, 1995: 279-299.
67. **Pittaway DE, Maxson WS, Wentz AC.** Spironolactone in combination drug therapy for unresponsive hirsutism. *Fertil Steril.* 1985;43:878-882.
68. **O'Brien RC, Cooper ME, Murray RM, Seeman E, Thomas AK, Jerums G.** Comparison of sequential cyproterone acetate/estrogen versus spironolactone/oral contraceptive in the treatment of hirsutism. *J Clin Endocrinol Metab.* 1991;72:1008-1013.
69. **Moghetti P, Castello R, Zamberlan N, et al.** Spironolactone, but not flutamide, administration prevents bone loss in hyperandrogenic women treated with gonadotropin-releasing hormone agonist. *J Clin Endocrinol Metab.* 1999;84:1250-1254.
70. **Castello R, Tosi F, Perrone F, Negri C, Muggeo M, Moghetti P.** Outcome of long-term treatment with the 5 alpha-reductase inhibitor finasteride in idiopathic hirsutism: clinical and hormonal effects during a 1-year course of therapy and 1-year follow-up. *Fertil Steril.* 1996;66:734-740.
71. **Ciotta L, Cianci A, Calogero AE, et al.** Clinical and endocrine effects of finasteride, a 5 alpha-reductase inhibitor, in women with idiopathic hirsutism. *Fertil Steril.* 1995;64:299-306.
72. **Velazquez EM, Mendoza S, Hamer T, Sosa F, Glueck CJ.** Metformin therapy in polycystic ovary syndrome reduces hyperinsulinemia, insulin resistance, hyperandrogenemia, and systolic blood pressure, while facilitating normal menses and pregnancy. *Metabolism.* 1994;43:647-654.
73. **Nestler JE.** Insulin and ovarian androgen excess. In: Azziz R, Nestler JE, Dewailly D, eds. *Androgen Excess Disorders in Women.* Philadelphia: Lippincott-Raven, 1997:473-483.
74. **Nestler JE, Jakubowicz DJ, Evans WS, Pasquali R.** Effects of metformin on spontaneous and clomiphene-induced ovulation in the polycystic ovary syndrome. *N Engl J Med.* 1998;338:1876-1880.
75. **Morin-Papunen LC, Koivunen RM, Ruokonen A, Martikainen HK.** Metformin therapy improves the menstrual pattern with minimal endocrine and metabolic effects in women with polycystic ovary syndrome. *Fertil Steril.* 1998;69:691-696.

76. **Moghetti P, Castello R, Negri C, et al.** Metformin effects on clinical features, endocrine and metabolic profiles, and insulin sensitivity in polycystic ovary syndrome: a randomized, double-blind, placebo-controlled 6-month trial, followed by open, long-term clinical evaluation. *J Clin Endocrinol Metab.* 2000;85:139-146.
77. **Thomsen HS, Morcos SK (ESUR Contrast Media Safety Committee).** Contrast media and metformin: guidelines to diminish the risk of lactic acidosis in non-insulin-dependent diabetics after administration of contrast media. *Eur Radiol.* 1999;9:738-740.
78. **Steingold KA, Lobo RA, Judd HL, Lu JK, Chang RJ.** The effect of bromocriptine on gonadotropin and steroid secretion in polycystic ovarian disease. *J Clin Endocrinol Metab.* 1986;62:1048-1051.
79. **Buvat J, Buvat-Herbaut M, Marcolin G, et al.** A double blind controlled study of the hormonal and clinical effects of bromocriptine in the polycystic ovary syndrome. *J Clin Endocrinol Metab.* 1986;63:119-124.
80. **Lemay A, Faure N.** Sequential estrogen-progestin addition to gonadotropin-releasing hormone agonist suppression for the chronic treatment of ovarian hyperandrogenism: a pilot study. *J Clin Endocrinol Metab.* 1994;79: 1716-1722.
81. **Halikias I, Lytras A, Syriou V, Tolis G.** Combined oral contraceptives and gonadotropin releasing hormone agonistic analogs in polycystic ovary syndrome: clinical and experimental studies. *Eur J Contracept Reprod Health Care.* 1997;2:213-224.
82. **Polson DW, Mason HD, Franks S.** Bromocriptine treatment of women with clomiphene-resistant polycystic ovary syndrome. *Clin Endocrinol (Oxf).* 1987;26:197-203.
83. **Futterweit W.** An endocrine approach to obesity. In: Simopoulos AP, Vanltallie TB, Gullo SP, Futterweit W, eds. *Obesity: New Directions in Assessment and Management.* New York: Charles Press, 1994: 96-121.
84. **Kiddy DS, Hamilton-Fairley D, Bush A, et al.** Improvement in endocrine and ovarian function during dietary treatment of obese women with polycystic ovary syndrome. *Clin Endocrinol (Oxf).* 1992;36:105-111.
85. **Pasquali R, Antenucci D, Casimirri F, et al.** Clinical and hormonal characteristics of obese amenorrheic hyperandrogenic women before and after weight loss. *J Clin Endocrinol Metab.* 1989;68:173-179.
86. **Jakubowicz DJ, Nestler JE.** 17 Alpha-hydroxyprogesterone responses to leuprolide and serum androgens in obese women with and without polycystic ovary syndrome after dietary weight loss. *J Clin Endocrinol Metab.* 1997; 82:556-560.

EXHIBIT R

Azziz et al., J. Clin. Endocrinol. Metab. 2004, 89(2):453-62

0021-972X/04/$15.00/0
Printed in U.S.A.

The Journal of Clinical Endocrinology & Metabolism 89(2):453–462

doi: 10.1210/jc.2003-031122

EXTENSIVE PERSONAL EXPERIENCE

Androgen Excess in Women: Experience with Over 1000 Consecutive Patients

R. AZZIZ, L. A. SANCHEZ, E. S. KNOCHENHAUER, C. MORAN, J. LAZENBY, K. C. STEPHENS, K. TAYLOR, AND L. R. BOOTS

Departments of Obstetrics and Gynecology (R.A., L.A.S., E.S.K., C.M., J.L., K.C.S., K.T., L.R.B.) and Medicine (R.A.), University of Alabama at Birmingham, Birmingham, Alabama 35233; Caracas Fertility Center (L.A.S.), Caracas, Federal District, Venezuela 1050; and the Health Research Council (C.M.), Mexican Institute of Social Security, Federal District, Mexico City, Mexico

The objective of the present study was to estimate the prevalence of the different pathological conditions causing clinically evident androgen excess and to document the degree of long-term success of suppressive and/or antiandrogen hormonal therapy in a large consecutive population of patients.

All patients presenting for evaluation of symptoms potentially related to androgen excess between October 1987 and June 2002 were evaluated, and the data were maintained prospectively in a computerized database. For the assessment of therapeutic response, a retrospective review of the medical chart was performed, after the exclusion of those patients seeking fertility therapy only, or with inadequate follow-up or poor compliance.

A total of 1281 consecutive patients were seen during the study period. Excluded from analysis were 408 patients in whom we were unable to evaluate hormonal status, determine ovulatory status, or find any evidence of androgen excess. In the remaining population of 873 patients, the unbiased prevalence of androgen-secreting neoplasms was 0.2%, 21-hydroxylase-deficient classic adrenal hyperplasia (CAH) was 0.6%, 21-hydroxylase-deficient nonclassic adrenal hyperplasia (NCAH) was 1.6%, hyperandrogenic insulin-resistant acanthosis nigricans (HAIRAN) syndrome was 3.1%, idiopathic hirsutism was 4.7%, and polycystic ovary syndrome (PCOS) was 82.0%. Fifty-nine (6.75%) patients had elevated androgen levels and hirsutism but normal ovulation. A total of 257 patients were included in the assessment of the response to hormonal therapy. The mean duration of follow-up was 33.5 months (range, 6–155). Hirsutism improved in 86%, menstrual dysfunction in 80%, acne in 81%, and hair loss in 33% of patients. The major side effects noted were irregular vaginal bleeding (16.1%), nausea (13.0%), and headaches (12.6%); only 36.6% of patients never complained of side effects.

In this large study of consecutive patients presenting with clinically evident androgen excess, specific identifiable disorders (NCAH, CAH, HAIRAN syndrome, and androgen-secreting neoplasms) were observed in approximately 7% of subjects, whereas functional androgen excess, principally PCOS, was observed in the remainder. Hirsutism, menstrual dysfunction, or acne, but not alopecia, improved in the majority of patients treated with a combination suppressive therapy; although more than 60% experienced side effects. (*J Clin Endocrinol Metab* 89: 453–462, 2004)

ANDROGEN EXCESS IS one of the most common endocrine disorders of reproductive-aged women, affecting approximately 7% of this population (1–3). Androgen excess results in the development of androgenic features in the women affected, with the development of hirsutism, androgenic alopecia, acne, ovulatory dysfunction, and, if extreme and prolonged, even virilization and masculinization. Disorders that result in androgen excess include specific identifiable disorders (*i.e.* disorders of inclusion) such as non-classic adrenal hyperplasia (NCAH), hyperandrogenic insulin-resistant acanthosis nigricans (HAIRAN) syndrome, and androgen-secreting neoplasms (ASNs). Alternatively, a number of androgen excess disorders are diagnosed by exclusion, such as the polycystic ovary syndrome (PCOS) and idiopathic hirsutism (IH), termed disorders of functional androgen excess (FAE).

Notwithstanding their clinical importance, the prevalence of the different pathological conditions causing or associated with androgen excess remains unclear. Although the most recognizable clinical feature of androgen excess may be hirsutism, it should be noted that not all patients with hirsutism have overt evidence of androgen excess, with some women suffering from what we understand to be IH (4). Alternatively, not all patients with an androgen excess disorder have hirsutism, as in the Asian patient with PCOS (5). However, in most studies, hirsutism has been used as a surrogate marker for androgen excess, with the largest study to date of

Abbreviations: ASN, Androgen-secreting neoplasm; BMI, body mass index; BTB, breakthrough bleeding; CAH, classic adrenal hyperplasia; DHEAS, dehydroepiandrosterone sulfate; FAE, functional androgen excess; GLU, glucose; HA, hyperandrogenemia; HAIRAN, hyperandrogenic insulin-resistant acanthosis nigricans; HOMA, homeostatic model assessment method; HOMA-%β-cell, percentage of β-cell function calculated by HOMA; HOMA-IR, insulin resistance calculated by HOMA; 17-HP, 17-hydroxyprogesterone; IH, idiopathic hirsutism; INS, insulin; mF-G, modified Ferriman-Gallwey; NCAH, non-classic adrenal hyperplasia; OC, oral contraceptive; 21-OH, 21-hydroxylase; PCOS, polycystic ovary syndrome; P4, progesterone; SPA, spironolactone; T, testosterone; WHR, waist to hip ratio.

JCEM is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.

consecutive patients potentially affected by androgen excess including only subjects presenting with hirsutism and/or androgenic alopecia (6). This has made it difficult to accurately assess the frequency of the lesser common androgen excess disorders.

The response of hirsutism, ovulatory dysfunction, and other features of androgen excess to hormonal therapy is important to determine but has been studied primarily with the use of individual agents. However, combination therapy including oral contraceptives (OCs), antiandrogens, or metformin has been suggested to be superior to monotherapy (7–9). Unfortunately, reports evaluating the results of these regimens have generally included 50 or fewer patients (7–13), limiting our assessment of the efficacy of this therapeutic regimen.

The objective of the present study was to report on our experience evaluating over 1000 consecutive patients consulting for symptoms potentially related to androgen excess. From this data, we have estimated the prevalence of the different pathological conditions causing clinically evident androgen excess and further documented the long-term success of suppressive and/or antiandrogen hormonal therapy in the treatment of these patients.

Subjects and Methods

Subjects

All patients presenting for the evaluation of symptoms potentially related to androgen excess to the reproductive endocrinology clinic (to R.A.) at the University of Alabama at Birmingham between October 1987 and June 2002 were included. The Institutional Review Board for Human Use at the University of Alabama at Birmingham approved the study. Patients evaluated included those presenting with oligo/amenorrhea, ovulatory dysfunction, excess hair growth, virilization, alopecia, or acne. The data were recorded and maintained prospectively in a computerized database (Alpha Four version 6.0; Alpha Software Corp., Burlington, MA).

Initial subject evaluation

All patients completed a uniform history form and underwent a complete physical examination. The following parameters were recorded prospectively: height, weight, race, age, gravidity, parity, degree of ovulatory/menstrual dysfunction, presence of acne, and hirsutism score. The body mass index (BMI) was calculated as kilograms per square meter. Beginning January 1993, the abdominal and hip circumferences were assessed as published previously (14), and the waist to hip ratio (WHR) was calculated. The presence of acne was recorded, but the severity was not generally scored. Excess body and facial terminal hair growth was assessed using a modified Ferriman-Gallwey (mF-G) hirsutism score (15).

The interval between bleeding episodes was assessed prospectively and classified as less than 26 d, 27–34 d, 35–44 d, 45 d to 3 months, and more than 3 months duration. This classification was established per the reported variations in menstrual cycle length observed by Treloar *et al.* (16). Patients with cycles greater than 35 d or less than 26 d were deemed to be oligo/anovulatory. In addition, beginning January 1995, ovulatory function was confirmed in all eumenorrheic (*i.e.* with a bleeding interval of 27–34 d) hirsute women by using the basal body temperature and measuring a cycle d 22–24 progesterone (P4) level. A level of P4 greater than 12.7 nmol/liter (4 ng/ml) was considered to represent ovulation, and eumenorrheic patients were then classified as "27–34 d plus ovulation" or "27–34 d plus oligo-ovulation."

In patients who had not received hormonal therapy for 3 months before their initial visit, the serum levels of total testosterone (T), free T, dehydroepiandrosterone sulfate (DHEAS), and 17-hydroxyprogesterone (17-HP) were obtained and recorded. Blood sampling for androgens was performed without regard to the time of the cycle or the day, although an effort was made to measure the 17-HP in the preovulatory phase of the menstrual cycle (17). Hyperandrogenemia was defined as an androgen value above the 95th percentile of 98 healthy control women [*i.e.* a total T $\geq$ 2.94 nmol/liter (88 ng/dl), free T $\geq$ 0.026 nmol/liter (0.75 ng/dl), or DHEAS $\geq$ 6.64 μmol/liter (2750 ng/ml)], as reported previously (1).

Patients with a basal level of 17-HP greater than 6.0 nmol/liter (2 ng/ml) underwent either a repeat 17-HP test or proceeded directly to an ACTH stimulation test, as described previously (17). If the repeat 17-HP was 6.0 nmol/liter or greater, patients underwent an acute ACTH stimulation test (see below). Patients with levels of total T above 8.67 nmol/liter (250 ng/dl) on at least two separate occasions, or who were deemed to demonstrate clinical features suggestive of an ASN (*e.g.* virilization), underwent a transvaginal sonogram and a computerized tomography scan of the adrenals at 5-mm intervals to exclude ovarian and adrenal neoplasms, respectively (18).

Patients with evidence of ovulatory dysfunction also underwent measurements of serum prolactin and thyroid-stimulating hormone levels to exclude a prolactin-secreting adenoma and thyroid dysfunction, respectively. If clinically suspected, screening for Cushing's syndrome was performed by either the overnight dexamethasone suppression test (*i.e.* the measurement of a cortisol level the morning after the administration of 1 mg dexamethasone orally at bedtime) through January 1990 or by measuring the 24-h urine-free cortisol content after that date.

In patients diagnosed with PCOS or HAIRAN syndrome (see below), fasting glucose (GLU) and insulin (INS) levels were obtained regularly after April 1996. Before those dates, basal and or GLU-stimulated INS and GLU levels were obtained selectively, primarily in patients clinically suspected of having INS resistance (*e.g.* by the presence of acanthosis nigricans on physical examination).

Acute ACTH stimulation test

When indicated (see above), the acute ACTH stimulation test was performed as described previously (17). In brief, all studies were started between 0730 and 0930 h in the fasting state and scheduled d 3–8 after a spontaneous vaginal bleed or after a withdrawal bleed induced by either 100 mg P4 in oil im or 300 mg/d oral micronized P4 for 7 d. Dexamethasone was not administered before the study. Three baseline samples were obtained 15 min apart and mixed (0-min sample). Immediately afterward, 0.25 mg ACTH-(1–24) (Cortrosyn; Organon Co., Orange, NJ) was administered iv over 60 sec, and blood was sampled 60 min later. Both the 0- and 60-min samples were assayed for 17-HP levels. If the stimulated 17-HP level was greater than 10 ng/ml, the patient was considered to have 21-hydroxylase (21-OH)-deficient NCAH (17). The diagnosis of NCAH was confirmed by the genotyping of CYP21, as described previously (17).

Differential diagnosis

Based on the above evaluation, two distinct types of androgen excess disorders, those with specific diagnoses ascertained by inclusion and those diagnosed by exclusion (defined as FAE disorders), were identified:

Specific diagnoses ascertained by inclusion testing: 1) ASNs, diagnosed by resection and histopathology of the tumor; 2) 21-OH CAH, diagnosed by review of prior records indicating evidence of congenital virilization in the presence of dramatically elevated basal levels of 17-HP, generally 90 nmol/liter (30 ng/ml) or greater (19); 3) 21-OH-deficient NCAH, diagnosed by an ACTH-stimulated 17-HP level greater than 10 ng/ml and confirmed by the genotyping of CYP21, as described previously (17); 4) HAIRAN syndrome, diagnosed by a fasting basal INS greater than 80 μIU/ml and/or an INS level during a 3-h oral glucose tolerance test greater than 300 μIU/ml, as described previously (20, 21).

Disorders of FAE, diagnosed by exclusion: 1) PCOS, diagnosed by the presence of ovulatory dysfunction in association with hirsutism and/or elevated androgen levels, after exclusion of the above specific disorders, as recommended by a 1990 National Institutes of Health/National Institute of Child Health and Human Development-sponsored conference on the subject (22); 2) IH, diagnosed by the presence of hirsutism accompanied by normal ovulation and normal androgens levels, after

exclusion of the above specific disorders (23); and 3) hyperandrogenemia (HA) plus hirsutism, comprised of individuals with elevated androgen levels and hirsutism but normal ovulation.

Assessment of response to suppressive hormonal therapy

To determine the response to suppressive hormonal therapy in patients with androgen excess, we retrospectively reviewed the charts of all androgen excess patients seen initially between October 1987 and June 2001. Using a uniform form, the documented treatment outcome and side effects as assessed by the patient were recorded. Patients included were those who had PCOS, HAIRAN, IH, or HA plus hirsutism (see above). We excluded those patients: 1) seeking infertility therapy; 2) with inadequate follow-up (*i.e.* duration of the follow-up visits was less than 6 months, or who did not return after their initial visit); or 3) with poor compliance with treatment or follow-up (*i.e.* who were noncompliant in taking the medication according to prescribed directions, or who were seen for care at greater than 18-month increments).

All patients with menstrual or ovulatory dysfunction received OCs when possible. Patients with unwanted hair growth and evidence of excess facial or body terminal hair growth received spironolactone (SPA) (200 mg if mF-G $\geq$ 8; 100 mg/d if mF-G = 3–7) in combination with the OC, to minimize the risks of teratogenicity. SPA was rarely used alone, except in the occasional hirsute patient who had previously undergone a hysterectomy or tubal ligation. Other treatment regimens were occasionally used, including glucocorticoids, insulin sensitizers, GnRH analogs, flutamide, finasteride, and other estrogen-progestin combinations, alone or in combination; the majority of these were used as part of clinical trials (24–26).

When side effects were identified, they were managed as follows. First, if side effects attributable to SPA occurred, the dose of SPA was decreased. On occasion, the SPA was changed to flutamide or finasteride. Second, if persistent breakthrough bleeding (BTB) on the OC occurred, despite short-term supplementation with an oral estrogen, then the OC was changed to one containing 50 μg ethinyl estradiol plus 1 mg ethyndiol diacetate. Third, if other side effects attributable to the OC occurred (*e.g.* headaches or migraines, weight gain, increased breast size, *etc.*), then the OC was changed to one containing 20 μg ethinyl estradiol or discontinued altogether. Fourth, women who were hypertensive, smokers more than 35 yr of age, or those who had a previous thromboembolic event on OCs/hormones were considered for cyclic progestogen treatment only, potentially in combination with SPA.

Statistical analysis

INS resistance and β-cell function were calculated by the homeostatic model assessment method (HOMA-IR and HOMA-%β-cell, respectively), as described previously (27). Two-group comparison of continuous variables was performed using a two-sample *t* test with adjustment for nonconstancy of variance, when necessary. More than two group means were compared using the ANOVA with *post hoc* least squares means pairwise comparisons. All categorical end points were compared using the χ^2 statistic or Fisher's exact test, when appropriate.

Results

General features of patients excluded from the study

A total of 1281 consecutive patients were seen during the time period of the study. Of these, 408 (31.9%) subjects were excluded from analysis (Table 1). In 157 (12.3% of total or 38.5% of excluded) patients, we were unable to evaluate hormonal status generally because they were already on hormonal therapy and were unwilling to discontinue the medication for the evaluation. In 113 (8.8% of total or 27.7% of excluded) patients, we were unable to confirm ovulatory status, and 138 (10.8% of total or 33.8% of excluded) women did not have evidence of androgen excess. In total, 873 subjects were included in the study.

Patients who were excluded from the study were slightly older (29.6 ± 15.4 *vs.* 27.7 ± 8.1 yr; $P < 0.02$) and of lesser BMI (29.1 ± 8.9 *vs.* 32.9 ± 9.6 kg/m^2; $P < 0.001$, respectively) than individuals who were included, although the differences were deemed to be of limited clinical significance. There was no difference in racial composition between the two groups, with subjects that were excluded being 14.3% black, 84.3% white, and 1.4% of other races compared with 15.0% black, 82.6% white, and 2.4% other races among those patients who were included.

TABLE 1. Patients excluded from study

Exclusion	Number	% of total excluded patients	% of total patients
Unable to evaluate hormonal status		38.5	12.2
Initial treatment	156		
Missing androgen levels	1		
Unable to determine ovulatory status		27.7	8.8
Postmenopausal[a]	19		
Premenarchal	4		
Hysterectomized[b]	4		
Vaginal agenesis	1		
Missing ovulatory assessment[c]	85		
Without androgen excess		33.8	10.8
Normal: 28	28		
Only HA: 26	26		
Only oligo: 84	84		
Total	408	100	32.8

[a] Except for one patient with an ASN.
[b] Hysterectomized, and without information regarding menstrual or ovulatory function.
[c] All with regular menses, d 26–34 in length, but without having ovulation verified by luteal progesterone level.

Clinical features of androgen excess patients included in the study

Of the 873 patients included in the study 5.1, 16.0, 21.8, 23.7, 17.4, 11.1, 3.2, and 1.6% were 15 yr old or younger, 16–20, 21–25, 26–30, 31–35, 36–40, 41–45, and 46 yr old or older, respectively. Clinically, 75.5% had hirsutism and 14.2% had acne, whereas 29.7% complained of infertility. Patients with hirsutism or infertility were slightly older [28.2 ± 7.6 *vs.* 26.5 ± 7.1 yr ($P < 0.005$) and 29.4 ± 5.6 *vs.* 27.1 ± 8.4 yr ($P < 0.001$), respectively] than women without these complaints. Patients with infertility were also more obese than their non-infertile counterparts (34.0 ± 8.9 *vs.* 32.5 ± 9.2 kg/m^2, $P < 0.03$, respectively). Alternatively, women with acne were younger and less overweight than those patients than non-acneic patients [23.6 ± 7.5 *vs.* 28.0 ± 7.0 yr ($P < 0.001$) and 29.5 ± 8.6 *vs.* 33.5 ± 9.1 kg/m^2 ($P < 0.001$)].

Oligo-ovulation was present in 770 (88.2%) patients; 86.4% of oligo-ovulatory patients had overt menstrual dysfunction (12 with polymenorrhea and the remainder with oligomenorrhea). One hundred five (13.6%) oligo-ovulatory patients had apparent eumenorrhea until evaluated more closely. Of hirsute women, 35 (5.3%) had apparent eumenorrhea, of which 14 (40%) had ovulatory dysfunction when evaluated by a luteal P4 level. Of the 124 patients with acne, 33.9% (or 4.8% of the total) had acne only, without hirsutism, although all these patients had both hyperandrogenemia and oligo-ovulation.

Overall, 58.2% of patients included were obese (BMI, ≥30.0 kg/m^2), with BMIs as follows: 1.8% were underweight (<19.0 kg/m^2), 20.6% were of normal weight (19.0–24.9 kg/m^2), 19.4% were preobese (25.0–29.9 kg/m^2), 18.7% were mildly obese (30.0–34.9 kg/m^2), 17.6% were moderately obese (35.0–39.9 kg/m^2), and 21.9% were severely obese (≥40 kg/m^2), according to 1998 World Health Organization and 1999 National Center for Health Statistics/Centers for Disease Control and Prevention criteria (28, 29).

Prevalence of abnormally elevated androgen measures

Overall, 77.8% of patients included in the analysis had hyperandrogenemia. Total T was elevated in 325 (37.9%), free T in 476 (55.5%), and DHEAS in 347 (40.4%) of the 858 subjects included. The prevalence of the different combinations of abnormal androgen levels is depicted in Table 2. Note that approximately one fifth of patients had no overt elevation in androgen levels, whereas another 10% had elevations in all three androgens evaluated.

Prevalence of the different diagnostic groups

Among the 873 patients entered into the study, five were on thyroid replacement for hypothyroidism at the time of their initial visit, and one additional patient was diagnosed with hypothyroidism at her evaluation (total prevalence of thyroid patients was diagnosed as such during her evaluation, for a total dysfunction of 0.7%). Two patients were receiving bromocriptine for a previous diagnosis of hyperprolactinemia, and one additional patient was diagnosed as such during her evaluation (for a total prevalence of hyperprolactinemia of 0.3%).

Of the 873 patients included in the study (Table 3), 59 (6.76%) had specific identifiable disorders including ASNs, 21-OH-deficient CAH, 21-OH-deficient NCAH, and HAIRAN syndrome diagnosed in 0.23, 0.69, 2.06, and 3.78% of the total population, respectively. In the remainder (814 or 93.24% of the total), disorders of FAE (*i.e.* of exclusion) were identified including PCOS, IH, and HA plus hirsutism in 82.02, 4.47, and 6.75% of patients, respectively.

Because ovulatory function was determined in all eumenorrheic hirsute women beginning January 1995, the unbiased incidence of IH was calculated as 4.68% (23 of 491 subjects studied). After June 1996, the fasting INS levels were routinely obtained, resulting in an unbiased incidence of HAIRAN syndrome of 3.12% (12 of 385). Finally, because three of the NCAH patients were diagnosed because they were relatives of a patient already diagnosed with this disorder and one additional patient was diagnosed before the study, the unbiased incidence of NCAH was 1.60% (14 of 873). Only two of our NCAH patients did not have a detectable mutation on genotyping, and both of these women had ACTH-stimulated 17-HP levels greater than 20 ng/ml, suggesting that their diagnosis was not in question.

TABLE 2. Prevalence of abnormally elevated androgen measures in 858[a] patients with androgen excess

Abnormal androgen(s)[b]	Number	Prevalence (%)
None	190	22.2
Total T only	20	2.4
Free T only	125	14.6
DHEAS only	146	17.0
Total T and free T	174	20.3
Total T and DHEAS	24	2.8
Free T and DHEAS	69	8.0
Total T, free T, and DHEAS	109	12.7
Total	858	100.0

[a] Note that 15 subjects did not have all three androgen results (total and free T and DHEAS) available and, thus, were not included in this analysis.

[b] Abnormal androgen levels were defined as an androgen value above the 95th percentile of 98 healthy control women [*i.e.* a total T ≥ 2.94 nmol/liter (88 ng/dl), free T ≥ 0.026 nmol/liter (0.66 ng/dl), or DHEAS ≥ 6.64 μmol/liter (2750 ng/ml)], as reported previously (1).

TABLE 3. Differential diagnosis of 873 patients evaluated for androgen excess

Diagnosis	Total no.	% Prevalence	% Unbiased prevalence[a]
Specific disorders			
ASN	2	0.23	
CAH	6	0.69	
NCAH	18	2.06	1.60
HAIRAN	33	3.78	3.12
Disorders of exclusion			
PCOS	716	82.02	
IH	39	4.47	4.68
HA + hirsutism	59	6.75	
Total	873	100.00	

[a] The unbiased prevalence of IH was calculated using the population screened after January 1995, because ovulatory function was determined in all eumenorrheic hirsute women beginning with that date. The unbiased prevalence of HAIRAN syndrome was calculated using the population screened after June 1996, when the fasting insulin levels were routinely obtained. The unbiased prevalence of NCAH was based on the number of unrelated patients diagnosed during the study period and who had not been identified *a priori* by either a previous diagnosis or by having affected relatives.

Comparison of diagnostic groups

In comparing clinical and laboratory features between diagnostic groups, we again excluded patients with CAH and ASN because of their small numbers. We observed (Table 4) that there was significant racial difference between the different diagnoses, with NCAH patients being exclusively of the white race, whereas a greater proportion of HAIRAN patients were black (36.4%).

HAIRAN syndrome patients were younger, and women with IH were older than all others. HAIRAN syndrome patients also had a greater body mass and WHR than all others. Although PCOS patients had a lower mean BMI and WHR than HAIRAN patients, they were, on average, larger and had a higher WHRs than the other diagnostic groups.

There was no difference between the groups regarding the prevalence of acne, whereas the differences observed in the frequency of oligo-ovulation and hirsutism were primarily related to the differences in the criteria for defining a diagnostic group (*e.g.* by definition, patients with IH and HA plus hirsutism have hirsutism with normo-ovulation). PCOS patients had the highest and HA plus hirsutism patients the lowest prevalence of infertility, respectively.

When comparing measures of INS action, including fasting glucose, INS, HOMA-IR, and HOMA-%β-cell, we analyzed only those values obtained after April 1996, when these

TABLE 4. Characteristics of the study population by diagnostic group

Variable	Diagnostic group				
	PCOS (n = 716)	HAIRAN (n = 33)	NCAH (n = 18)	IH (n = 39)	HA + hirsutism (n = 59)
Age (yr)	27.67 ± 7.26	23.10 ± 7.96[a]	28.28 ± 10.65	32.58 ± 9.29[a]	27.71 ± 8.99
BMI (kg/m^2)	33.39 ± 9.26[a]	38.69 ± 9.21[a]	29.00 ± 6.32	28.65 ± 7.13	28.65 ± 6.72
WHR	0.83 ± 0.09[f]	0.87 ± 0.092[e,f]	0.83 ± 0.069	0.81 ± 0.11	0.79 ± 0.08
Total T (ng/dl)[g]	90.28 ± 55.67[e]	111.91 ± 136.72[b,e,f]	116.24 ± 64.61[b,e,f]	57.40 ± 13.84[b,c,d]	75 ± 19.61[c,d]
SHBG (nmol/liter)[g]	182 ± 71[c,d,e]	155 ± 41[a]	235 ± 15[b,c,f]	243 ± 91[b,c,e]	186 ± 68[d,e]
Free T (ng/dl)[g]	0.92 ± 0.57[f]	0.95 ± 0.54	1.27 ± 0.55[b,f]	0.47 ± 0.15[a]	0.75 ± 0.27[b,e,f]
DHEAS (ng/ml)[g]	2350.8 ± 1289.6[a]	1897.6 ± 1215.9[b,d,f]	3485.6 ± 1691.2[b,c,e]	368.46 ± 722.18[b,d,f]	2910.8 ± 1204.9[b,c,e]
mF-G score	8.2 ± 4.9[e,f]	10.1 ± 5.22[b]	7.78 ± 3.98	9.0 ± 3.6	9.6 ± 3.2[b]
Fasting plasma INS (μU/ml)[h]	21.1 ± 1.0[c,e,f]	88.1 ± 4.5[a]	9.5 ± 8.4[c]	12.0 ± 4.5[b,c]	12.9 ± 3.6[b,c]
Plasma GLU (mg/dl)[h]	89.5 ± 1.2	92.6 ± 5.4	91.5 ± 10.1	89.4 ± 5.4	86.5 ± 4.2
HOMA-IR (μU·mol/ml)[h]	4.70 ± 0.37	22.26 ± 1.60[a]	2.22 ± 3.39	2.78 ± 2.53	2.80 ± 3.06
HOMA-%β-cell (μU/mol)[h]	335.3 ± 369.3	193.7 ± 1012.3	113.0 ± 133.3	173.7 ± 216.2	213.9 ± 158.2
Race					
White (%)	83.2[c]	63.6[b,d,f]	100.0[c]	79.0	84.7[c]
Black (%)	14.4[c]	36.4[b,d,f]	0.0[c]	18.4	10.2[c]
Other (%)	2.4	0.0	0.0	2.6	5.0
Acne (%)	14.5	6.1	22.2	17.9	11.9
Oligo-ovulatory (%)[i]	100.0[d,e,f]	100.0[e,f]	88.9[b,e,f]	0.0[b,c,d]	0.0[b,c,d]
Infertile (%)	32.7[c,f]	13.8[b]	22.2	23.1[f]	6.8[b,e]
Hirsute (%)[i,j]	72.2[e,f]	84.8[e,f]	72.2[e,f]	100.0[b,c,d]	100.0[b,c,d]
Obesity (%)[k]	60.0[c,e,f]	90.0[b,d,e,f]	50.0[c]	35.9[b,c]	37.3[b,c]

The data are mean ± SD. Variables: total and free T [conversion factor to SI units (nmol/liter) is 0.03467]; DHEAS [conversion factor to SI units (μmol/liter) is 0.002714]; INS [conversion factor to SI units (pmol/liter) is 7.175]; GLU [conversion factor to SI units (mmol/liter) is 0.05551].

[a] Significantly different from all other diagnostic groups.
[b] Significantly different from PCOS.
[c] Significantly different from HAIRAN.
[d] Significantly different from NCAH.
[e] Significantly different from IH.
[f] Significantly different from HA + hirsutism.
[g] IH patients were excluded from the analysis of androgen measures.
[h] The values of INS, GLUC, HOMA-IR, and HOMA-%β-cell depicted and compared are for the period of April 1996 through June 2002, when these measures were obtained routinely.
[i] By definition, patients with IH or with HA + hirsutism are always hirsute and normo-ovulatory.
[j] Hirsutism is defined as a mF-G score of 6 or greater.
[k] Obesity is defined as 30 kg/m^2 or greater (28, 29).

measures were obtained routinely. As expected, HAIRAN patients had higher fasting INS and HOMA-IR values than all other subjects (Table 4). Patients with PCOS also had higher INS levels than women with IH or HA plus hirsutism. There were no differences in mean fasting GLU levels or HOMA-%β-cell values between groups.

Long-term response to hormonal suppressive therapy

Subjects. To assess the long-term response of androgen excess patients to hormonal suppressive therapy, we reviewed 736 consecutive patient charts. Of these, we excluded 102 patients (13.8% of the total) because we were unable to locate their full medical record. Of the remaining patients, 133 (18.1% of total or 21.0% of those patients in whom the chart was found) were excluded because they sought fertility therapy only. Another 244 (33.1% of total or 38.5% of those patients in whom the chart was found) patients were further excluded from analysis due to inadequate duration of follow-up (n = 172 or 23.4% of total), or for noncompliance with either treatment or follow-up (n = 72 or 9.8% of total). Patients who had inadequate duration of follow-up or were not compliant with either treatment or follow-up were more likely to be black but not different in age, BMI, or initial mF-G score than those subjects who were compliant with therapy and follow-up (Table 5). Overall, among patients with a chart available, 31.7% of black patients were compliant and had adequate follow-up compared with 54.1% of white patients.

Of the 257 patients who met the inclusion criteria for compliance and follow-up, 246 (95.7%) presented with an initial (although not sole) complaint of excess body and/or facial hair growth, 178 (69.3%) with some form of menstrual dysfunction, 42 (16.3%) with acne, 12 (4.7%) with hair loss, and 22 (8.6%) with weight gain or obesity. It should be noted that the patient's initial complaint did not necessarily coincide with findings on physical examination. For example, of 246 patients presenting with a complaint of hirsutism, 38 (15.4%) had a mF-G score 5 or less, whereas five women with a FG score of 6 or greater did not initially note excess hair growth.

Two hundred (77.8%) of the patients included in this analysis had PCOS, 27 (10.5%) had hirsutism and hyperandrogenemia but normal ovulatory function, 12 (4.7%) had HAIRAN syndrome, 10 (3.9%) had NCAH, and eight (3.1) had IH. Overall, the small numbers of subjects in some of these categories precluded a separate analysis of therapeutic success rates by diagnostic group. The initial treatment prescribed was OC plus SPA in 181 (70.4%) patients, OC only in 22 (8.6%) patients, SPA only in 30 (11.7%) patients, and other

TABLE 5. Characteristics of androgen excess patients who were included *vs.* those who were excluded due to an inadequate duration of follow-up (F/U), noncompliance with either treatment or follow-up, or inability to retrieve full medical record

Characteristics	Included[a] (n = 257)	<6 months F/U[b] (n = 172)	Noncompliant[b] (n = 72)	Chart not found[b] (n = 102)
Race[c]				
White	230 (90%)	140 (81%)	55 (77%)	84 (82.3%)
Black	20 (8%)	27 (16%)	16 (22%)	17 (16.7%)
Other	7 (3%)	5 (3%)	1 (1%)	1 (1%)
Age (yr)				
Mean ± SD	27.6 ± 8.9	28.2 ± 8.1	26.2 ± 8.2	28.2 ± 7.5
Range	13–48	13–55	14–47	12–51
BMI (kg/m^2)				
Mean ± SD	31.8 ± 9.5	33.8 ± 9.2	33.3 ± 9.7	31.6 ± 8.4
Range	17.7–65.0	16.5–67.4	17.1–65.5	18.4–54.8
Initial mF-G score				
Mean ± SD	9.3 ± 4.7	8.9 ± 4.9	9.6 ± 4.3	8.1 ± 5.0
Range	0–27	0–31	2–18	0–24

[a] Patients who were compliant with both treatment and F/U were included in the analysis.
[b] Patients with inadequate F/U, noncompliance with either treatment or F/U, in whom we were unable to retrieve their full medical record, were excluded from the analysis.
[c] Significant difference in race between the groups ($\chi^2 = 15.743$; $P < 0.005$).

TABLE 6. Self-reported treatment outcome in hyperandrogenic patients receiving hormonal suppression and who were compliant with therapy

Clinical feature	Treatment outcome	No. of patient responses	Overall % response
Excess body and/or hair growth	Better	207	85.9
	Same	30	12.4
	Worse	4	1.7
Menstrual dysfunction	Better	146	79.8
	Same	26	14.2
	Worse	11	6.0
Acne	Better	29	80.6
	Same	6	16.7
	Worse	1	2.8
Scalp hair loss	Better	3	33.3
	Same	2	22.2
	Worse	4	44.0

TABLE 7. Prevalence of side effects among hyperandrogenic patients receiving suppressive therapy and who were compliant with their treatment

Side effect	No. of patients with side effect	% of patients with side effect
Irregular vaginal bleeding or BTB	41	16.1
Nausea or vomiting	33	13.0
Headache or migraines	32	12.6
Depression or mood changes	30	11.8
Vaginal infections/irritation	13	5.1
Fatigue	13	5.1
Muscle cramps	8	3.1
Decreased libido	7	2.8
Dyspepsia or heartburn	7	2.8
Abdominal pain	6	2.4
Weight gain	6	2.4
Hot flashes	5	2.0
Insomnia	4	1.6
Alopecia	3	1.2
Weight loss	2	0.8
Acne	2	0.8
Night sweats	2	0.85
Dry eyes	2	0.8
Bloating	2	0.8
Heart palpitations	2	0.8
No symptoms	93	36.6

medications or combinations in 24 (9.3%) patients. The mean follow-up period was 33.5 ± 30.1 months (range, 6–129).

Outcome. Self-reported treatment outcomes are depicted in Table 6. The success rate (*i.e.* patients reporting that they were better) in women who were compliant with therapy ranged from 80–86% for hirsutism, menstrual dysfunction, and acne. The success rate of treating alopecia, albeit with a small number of patients with this complaint, was significantly lower at 33%. There were no differences in mean BMI or age between patients experiencing improvement in either hirsutism or menstrual dysfunction. A meaningful analysis could not be performed for acne or alopecia due to the small number of subjects in each group.

The initial and follow-up mF-G scores were available in 176 subjects of a total of 213 patients presenting with hirsutism (*i.e.* mF-G score ≥ 6). The mean initial mF-G score was 10.5 ± 4.1, and the last recorded mF-G score was 4.7 ± 3.4, with a mean net change in mF-G score of 5.9 ± 4.1 in 3.4 ± 2.7 yr. The mean rate of decrease in the mF-G score was −2.8 ± 3.3 per year. Overall, 65.9% of patients had a mF-G score of 5 or less at their last recorded follow-up evaluation. If only those patients who received combination suppressive therapy (OC plus SPA) were analyzed, the mean net change in mF-G score was −6.2 ± 4.1, with a self reported rate of improvement of 89.0%. One hundred one (47.4% of the total) patients with a mF-G score or 6 or greater reported using electrology. Patients who used electrology concomitantly with hormonal suppression had a greater net change in their mF-G score compared with those hirsute women not using electrology (3.0 ± 3.0 *vs.* −2.6 ± 2.4; $P < 0.05$).

Side effects. The most common side effects of hormonal suppressive therapy reported were headache, nausea, and irregular vaginal bleeding (Table 7). Only 93 (36.6%) subjects assessed did not report any side effects. Of patients having BTB, 38.2% stated that the hormonal therapy did not improve or actually worsened their menstrual dysfunction compared with 16.1% of patients without BTB ($P < 0.005$).

Discussion

We have reported on our experience evaluating over 800 patients with androgen excess. Of the patients included in the study, two thirds were 30 yr old or younger, with 20% of patients under the age of 20 yr at the time of their initial visit. Assuming that most androgen excess disorders begin to become clinically evident peripubertally, these data suggest that up to 80% of patients are not recognized, evaluated, and treated in a timely fashion. Approximately 60% of patients studied were obese (BMI, $\geq$30.0 kg/m^2), and 20% were severely obese (BMI, $\geq$40 kg/m^2). Hence, the prevalence of obesity among androgen excess patients in this study is at least 2-fold higher than that of the general population. For example, data from the National Health and Nutrition Examination Survey project indicate that the prevalence of obesity in women 20–34 yr of age of all ethnicities/races was 25.8% in the 1999–2000 survey (30, 31).

In our population, clinically evident androgen excess, namely hirsutism, was observed in 75% of subjects. We should note that the prevalence of hirsutism would have been lower if the proportion of patients of Asian descent (<1% in the studied population) were greater, because these individuals tend to have lower degrees of excess hair growth (5). An improvement in hirsutism with combination hormonal suppression (an OC and SPN) was reported by approximately 90% of patients, and approximately 65% of these women had minimal excess hair growth (*i.e.* a mF-G score $\leq$5) at their last recorded follow-up examination after a mean treatment period of approximately 3.5 yr. This success rate is consistent with previous reports evaluating antiandrogen therapy, alone or in combination with OCs, for the treatment of hirsutism (32, 33). Our data also indicated that patients who used electrology in addition to hormonal suppression had a greater decrease in the hair growth than those who did not, suggesting that this hair destruction technique should be encouraged as an adjuvant in the management of the hirsute woman.

Acne was the sole complaint or finding in 4.8% of hyperandrogenic patients evaluated, all of whom were both hyperandrogenemic and oligo-ovulatory. We and others have noted previously that the prevalence (34–36) of hyperandrogenemia among acneic patients is significant and that this clinical feature may be used as a marker for androgen excess, regardless of the patient's age. However, we should note that our study design does not allow us to estimate the prevalence of androgen excess among acneic patients, because we included only hyperandrogenic patients into our study. The self-reported improvement in acne with hormonal suppression was approximately 80%. Similar success rates have been reported in randomized trials of acneic women treated with an OCP only (37–40). Although spironolactone has also been found to have a positive effect on acne, a systematic review of published data could not determine the effectiveness of treatment and its value in clinical practice due to the small sample populations involved in previous trials (32).

Scalp hair loss was a complaint in 14 (4.0%) patients. All but one had ovulatory dysfunction, and all were also hirsute. Our data are consistent with that of others, indicating that the prevalence of androgen excess is high in alopetic women, particularly in those with ovulatory dysfunction (41). As for acne, our study design does not allow us to estimate the prevalence of androgen excess among alopetic patients, in general. Hormonal suppression, even in these hyperandrogenic patients, had limited effect on their alopecia, although the response of hirsutism and menstrual dysfunction in these patients was overall similar to that of the study population as a whole. The poor response of androgenic alopecia to androgen blockade and/or hormonal suppression is consistent with the results of a recent randomized trial (42).

Almost 90% of our patients had ovulatory dysfunction, approximately 85% with obvious oligomenorrhea and approximately 1.5% with polymenorrhea (cycles less than 26 d in length). Almost 15% of oligo-ovulatory patients had apparent eumenorrhea until evaluated more closely using a luteal (d 22–24) phase P4 level with or without basal body temperature monitoring. Of hirsute eumenorrheic women (5.3% of the total hirsute population), 40% were actually found to have ovulatory dysfunction when evaluated more closely. These data confirm the high prevalence of ovulatory dysfunction among hirsute women with regular menstrual cycles reported by others and ourselves (23, 43). Thus, direct assessment of ovulatory function should be a routine part of the evaluation of the hirsute woman with apparently regular episodes of vaginal bleeding. The success of hormonal therapy in subjectively improving menstrual dysfunction was approximately 80%. The high rate of BTB (~16%) may have precluded a higher success rate in the treatment of this clinical feature, because patients who experienced BTB were twice as likely as those who did not to state that hormonal therapy did not improve their menstrual dysfunction (38.2 *vs.* 16.1%).

Overall, approximately 78% of patients included in this study had detectable hyperandrogenemia, indicating that approximately 30% of patients did not have overt evidence of hyperandrogenemia, consistent with previous reports (1, 3, 44, 45). Free T was the most prevalent marker, solely elevated in 15% of patients and in combination with abnormalities of either total T or DHEAS in another approximate 40%. DHEAS also proved to be useful because it was the only laboratory abnormality detected in 17% of patients studied. However, total T was much less useful, the only androgen to be elevated in only 2% of subjects.

Because approximately 25% of our patients are nonhirsute and primarily diagnosed by having laboratory evidence of hyperandrogenemia, our data indicate that the measurement of free T and DHEAS are indicated in the evaluation of the nonhirsute woman suspected of suffering from androgen excess. Nonetheless, a few caveats are evident. First, the diagnostic value of free T is closely related to the assay method used. Recommended methods for the assessment of free T includes equilibrium dialysis (46, 47), calculation of free T from the measurement of SHBG and total T (48, 49), or ammonium sulfate precipitation (50). In general, current direct assays for the measurement of free T have limited value, particularly in the evaluation of the hyperandrogenic woman (46, 47). Second, the impact of the age-related decline in DHEAS levels, suggesting the need for age-related nor-

mative values, on the diagnostic value of this metabolite remains unclear (51).

In addition to the evaluation of the patient with uncertain androgenization, some investigators feel that the measurement of total T and DHEAS has some value in the detection of ASNs (52). However, more recent data suggest that the best predictor of these neoplasms is the clinical presentation, because androgen measures may be misleading and have a low positive predictive value (16, 53–55). In the population studied, both of our patients with ovarian androgen-secreting tumors had total T levels greater than 6.94 nmol/liter (200 ng/dl) when diagnosed. One patient was 65 yr old and had a rapid and sudden onset of severe hyperandrogenic signs and symptoms, suggesting the presence of an ASN, regardless of androgen levels (18). However, the second patient was originally mistaken for having 21-OH-deficient NCAH, and it was only after the circulating total T levels rose persistently to levels above 6.94 nmol/liter despite treatment with combination therapy that the correct diagnosis was suspected. Hence, although the positive predictive value of a total T level greater than 6.94 nmol/liter (200 ng/dl) is approximately 10% (18), this androgen may still have some value in identifying less clinically obvious patients with an ASN.

Thyroid dysfunction was found to be a relatively uncommon abnormality in our hyperandrogenic patients. Ferriman and Purdie (56) reported that none of their 467 hirsute women had a thyroid abnormality, whereas we diagnosed only one of our 873 patients as having hypothyroidism (although five others were already on thyroid replacement before their evaluation). This prevalence is similar or actually less than that reported by other investigators in the general population of women of similar age (57–61). Hyperprolactinemia was also rare, diagnosed in only one (0.3%) of our patients during their evaluation, a prevalence similar to that found by other investigators (6, 55, 62–64).

Approximately 7% of our androgen excess patients suffered from specific and definable disorders of inclusion, including ASNs (0.2%), CAH (0.7%), 21-OH-deficient NCAH (1.6%), and the HAIRAN syndrome (3.1%). Whereas ASNs are diagnosed by pathological examination and 21-OH-deficient NCAH by CYP21 genotyping, the diagnosis of HAIRAN syndrome is less clear. Although previous investigators have established diagnostic levels of basal or GLU-stimulated INS levels that are 3- to 5-fold the upper normal limits and some patients may have overt lipodystrophy (20, 21), it may be argued that many of these women simply represent a more metabolically affected group for PCOS. Only more intensive family and genetic studies will elucidate the true etiology and diagnostic criteria for this syndrome.

The remainder of patients evaluated had FAE or disorders of exclusion, including PCOS (82%), IH (4.7%), and hirsute and hyperandrogenemic patients with normal ovulatory function (6.8%). None of our patients were found to have Cushing's syndrome or adrenal neoplasms. Also, although a few patients were seen with minimal clinical evidence of androgen excess (*e.g.* acne and/or seborrhea) after the administration of an androgenic drug, none were found to have sufficiently significant clinical and/or biochemical hyperandrogenism to meet the entry criteria into this study.

Notwithstanding significant variations in how the disorders were defined and the ethnicity of the populations studied, our prevalences are generally similar to those of other investigators (6, 55, 62–64). In these studies, the prevalence of ASNs ranged from 0.6–2.1% (ovarian ASNs, 0.3–1%; adrenal ASNs, 0–2.1%), NCAH from 1–3%, drug-related hyperandrogenism from 0–1%, and Cushings syndrome from 0–1%. The widest variations were found in the prevalences of PCOS (37.8–78%) and IH (15–38.7%), although the relative degrees were in general agreement with those of our study. We should note that previous reports differ from the present study by using criteria for the disorders diagnosed that was ambiguous or is currently outdated, and for not adjusting the prevalence rates reported for the potential bias inherent to the screening methods used.

Comparing the various diagnostic groups, we noted that HAIRAN syndrome patients were younger, more obese, more likely to be black, more hirsute, more INS resistant, and had the lowest SHBG levels compared with the other phenotypes. Patients with PCOS had the second highest degree of obesity and INS resistance and were the most likely to be infertile. Patients with IH were, on average, older and, as expected, had lower total T, free T, and DHEAS levels. Women with NCAH had the highest mean DHEAS and total and free T levels and were exclusively of the white race.

Approximately 7% of androgen excess patients had hirsutism and hyperandrogenemia, but with normal ovulation, and represent a diagnostic conundrum. Compared with PCOS patients, these women had a similar mean age but lower mean BMI, were less likely to be infertile, reflecting their normal ovulatory status; and had lower free T levels despite being more hirsute. They did, however, have higher DHEAS levels than PCOS patients. On average, these patients were slightly younger and had higher DHEAS and free T levels compared with IH patients. However, they had similar mean BMI and obesity rates, measures of INS action, and mF-G scores as patients with IH. These data suggest that patients with hirsutism and hyperandrogenemia, but with normal ovulation, may more closely resemble IH than PCOS patients and may represent individuals with predominantly adrenal androgen excess. Whether this clinical presentation represents a stable phenotype or is a transition form to PCOS remains to be determined.

Although we have discussed the overall results of hormonal therapy above, a few facts are worth highlighting. First, of the 501 patients who were seeking hormonal suppression of their androgenic symptoms and for whom the chart was located, approximately 50% had either less than 6 months of follow-up or were noncompliant with either treatment or follow-up. These patients were more likely to be black but did not differ in the apparent severity of their disorder. Impacting on compliance may be the high rate of side effects with therapy. Although few, if any, serious side effects were observed, only 37% of subjects studied did not report any side effects. Notably, in Alabama, few other practitioner options were available for patients desiring management of their hyperandrogenism. These data suggest that we must identify strategies for improving the level of education and compliance of these women.

The present report has a number of strengths: the large number of patients included; the fact that baseline features

were obtained and maintained prospectively; the careful phenotyping implemented, including early recognition of the possibility of HAIRAN syndrome and NCAH and use of a standardized mF-G hirsutism scoring system; the recognition that diagnostic criteria evolved throughout the study; the uniformity of the assessment as performed by one investigator (R.A.); and the recognition that the investigator served a large catchment area. However, the study also has a few notable limitations: the fact that the protocol and definitions were modified over time, consistent with changes in our understanding of the disorders under consideration; the retrospective nature of our assessment of the presence of alopecia and of the therapeutic outcomes; and the increasing reputation of the examiner in this area of study throughout the time of the study, which may have increased the proportion of patients seen with rare causes of androgen excess. Nonetheless, it is unlikely that these limitations would significantly affect the results of the study.

In conclusion, in this large study of consecutive patients, PCOS was observed in approximately 80%, IH in approximately 5%, HAIRAN syndrome in approximately 3%, and NCAH in approximately 1.5% of patients, and ASNs were observed in approximately one of 500 androgen excess patients. The etiology of 7% of women who demonstrated hirsutism and hyperandrogenemia, but normal ovulation, remains to be determined. Over 80% of androgen excess patients compliant with their therapy experienced an improvement in hirsutism, menstrual dysfunction, and acne with a suppressive hormonal regimen, although only 36% were free of side effects. Alternatively, hair loss, although an infrequent complaint in our population, failed to improve in almost 70% of patients affected. Notably, almost 50% of patients did not appear to be compliant with therapy. These data indicate that PCOS is the most frequent cause of androgen excess, with all other remaining causes affecting a minority of patients, and that, with the exception of alopecia, therapeutic success is high in patients who are compliant with hormonal therapy.

Acknowledgments

Received July 1, 2003. Accepted October 31, 2003.

Address all correspondence and requests for reprints to: Ricardo Azziz, M.D., M.P.H., M.B.A., Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, 8635 West Third Street, Suite 160 W, Los Angeles, California 90048. E-mail: azzizr@cshs.org.

This work was supported in part by National Institutes of Health Grants RO1-HD29364 and K24-D01346 (to R.A.).

References

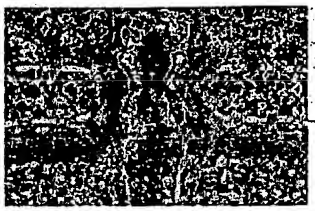
1. **Knochenhauer ES, Key TJ, Kahsar-Miller M, Waggoner W, Boots LR, Azziz R** 1998 Prevalence of the polycystic ovary syndrome in unselected black and white women of the southeastern United States: a prospective study. J Clin Endocrinol Metab 83:3078–3082
2. **Diamanti-Kandarakis E, Kouli CR, Bergiele AT, Filandra FA, Tsianateli TC, Spina GG, Zapanti ED, Bartzis MI** 1999 A survey of the polycystic ovary syndrome in the Greek island of Lesbos: hormonal and metabolic profile. J Clin Endocrinol Metab 84:4006–4011
3. **Asuncion M, Calvo RM, San Millan JL, Sancho J, Avila S, Escobar-Morreale HF** 2000 A prospective study of the prevalence of the polycystic ovary syndrome in unselected Caucasian women from Spain. J Clin Endocrinol Metab 85:2434–2438
4. **Azziz R, Carmina E, Sawaya ME** 2000 Idiopathic hirsutism. Endocr Rev 21:347–362
5. **Carmina E, Koyama T, Chang L, Stanczyk FZ, Lobo RA** 1992 Does ethnicity influence the prevalence of adrenal hyperandrogenism and insulin resistance in polycystic ovary syndrome. Am J Obstet Gynecol 167:1807–1812
6. **O'Driscoll JB, Mamtora H, Higginson J, Pollock A, Kane J, Anderson DC** 1994 A prospective study of the prevalence of clear-cut endocrine disorders and polycystic ovaries in 350 patients presenting with hirsutism or androgenic alopecia. Clin Endocrinol (Oxf) 41:231–236
7. **Pittaway DE, Maxson WS, Wentz AC** 1985 Spironolactone in combination drug therapy for unresponsive hirsutism. Fertil Steril 43:878–882
8. **Ibanez L, de Zegher F** 2003 Low-dose combination of flutamide, metformin and an oral contraceptive for non-obese, young women with polycystic ovary syndrome. Hum Reprod 18:57–60
9. **Elter K, Imir G, Durmusoglu F** 2002 Clinical, endocrine and metabolic effects of metformin added to ethinyl estradiol-cyproterone acetate in non-obese women with polycystic ovarian syndrome: a randomized controlled study. Hum Reprod 17:1729–1737
10. **Chapman MG, Dowsett M, Dewhurst CJ, Jeffcoate SL** 1985 Spironolactone in combination with an oral contraceptive: an alternative treatment for hirsutism. Br J Obstet Gynaecol 92:983–985
11. **Cusan L, Dupont A, Gomez JL, Tremblay RR, Labrie F** 1994 Comparison of flutamide and spironolactone in the treatment of hirsutism: a randomized controlled trial. Fertil Steril 61:281–287
12. **Erenus M, Yucelten D, Gurbuz O, Durmusoglu F, Pekin S** 1996 Comparison of spironolactone-oral contraceptive versus cyproterone acetate-estrogen regimens in the treatment of hirsutism. Fertil Steril 66:216–219
13. **Kelestimur F, Sahin Y** 1998 Comparison of Diane 35 and Diane 35 plus spironolactone in the treatment of hirsutism. Fertil Steril 69:66–69
14. **Steinkamp RC, Cohen NL, Gaffey WR, McKee YT, Bron G, Siri WE, Sargeant TW, Isaacs E** 1965 Measures of body fat and related factors in normal adults - II. J Chronic Dis 18:1291–1307
15. **Hatch R, Rosenfield RL, Kim MH, Tredway D** 1981 Hirsutism: implications, etiology, and management. Am J Obstet Gynecol 140:815–830
16. **Treloar AE, Boynton RE, Behn BG, Brown BW** 1970 Variation of the human menstrual cycle through reproductive life. Int J Fertil 12:77–126
17. **Azziz R, Hincapie LA, Knochenhauer ES, Dewailly D, Fox L, Boots LR** 1999 Screening for 21-hydroxylase deficient non-classic adrenal hyperplasia among hyperandrogenic women: a prospective study. Fertil Steril 72:915–925
18. **Waggoner W, Boots LR, Azziz R** 1999 Total testosterone and DHEAS levels a predictors of androgen-secreting neoplasms: a populational study. Gynecol Endocrinol 13:394–400
19. **Grumbach MM, Hughes I, Conte FA** 2002 Disorders of sexual differentiation. In: Larsen PR, Kronenberg HM, Melmed S, Polonsky KS, eds. Williams textbook of endocrionology. 10th ed. New York: Saunders; 842–1002
20. **Barbieri RL, Ryan KJ** 1983 Hyperandrogenism, insulin resistance, and acanthosis nigricans syndrome: a common endocrinopathy with distinct pathophysiologic features. Am J Obstet Gynecol 147:90–101
21. **Moller DE, Cohen O, Yamaguchi Y, Azziz R, Grigorescu F, Eberle A, Flier JS** 1994 Prevalence of mutations in the insulin receptor gene in subjects with features of the type A syndrome of insulin resistance. Diabetes 43:247–255
22. **Zawadzki JK, Dunaif A** 1990 Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. In: Dunaif A, Givens JR, Haseltine F, Merriam GR, eds. Polycystic ovary syndrome. Boston: Blackwell Scientific Publications; 377–384
23. **Azziz R, Waggoner WT, Ochoa T, Knochenhauer ES, Boots LR** 1998 Idiopathic hirsutism: an uncommon cause of hirsutism in Alabama. Fertil Steril 70:274–278
24. **Azziz R, Ochoa TM, Bradley Jr EL, Potter HD, Boots LR** 1995 Leuprolide and estrogen versus oral contraceptive pills for the treatment of hirsutism: a prospective randomized study. J Clin Endocrinol Metab 80:3406–3411
25. **Azziz R, Black VY, Hines GA, Boots LR** 1999 Ovulation after glucocorticoid suppression of adrenal androgens in the polycystic ovary syndrome is not predicted by the basal dehydroepiandrosterone sulfate level. J Clin Endocrinol Metab 84:946–950
26. **Azziz R, Ehrmann D, Legro RS, Whitcomb RW, Fereshetian AG, O'Keefe M, Ghazzi MN, for the PCOS/Troglitazone Study Group** 2001 Troglitazone improves ovulation and hirsutism in the polycystic ovary syndrome: a multicenter, double blind, placebo-controlled trial. J Clin Endocrinol Metab 86:1626–1632
27. **Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC** 1985 Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28:412–419
28. **World Health Organization** 1997 Obesity: preventing and managing the global epidemic. Report of a WHO Consultation of obesity. Geneva: World Health Organization
29. **National Center for Health Statistics** 2002 Table 70: Healthy weight, overweight, and obesity among persons 20 years of age and over, according to sex, age, race, and Hispanic origin: United States, 1960–62, 1971–74, 1976–80, 1988–94, and 1999–2000. In: Pastor P, Makuc DM, Reuben C, Xia H, eds. Health, United States, 2002 with chartbook on trends in the health of Americans. Hyattsville, MD: National Center for Health Statistics
30. **National Institutes of Health** 1998 Clinical guidelines on the identification,

evaluation, and treatment of overweight and obesity in adults. Bethesda, MD: National Institutes of Health, National Heart, Lung, and Blood Institute
31. **Flegal KM, Carroll MD, Ogden CL, Johnson CL** 2002 Prevalence and trends in obesity among US adults, 1999–2000. JAMA 288:1723–1727
32. **Farquhar C, Lee O, Toomath R, Jepson R** 2001 Spironolactone versus placebo or in combination with steroids for hirsutism and/or acne. Cochrane Database Syst Rev CD000194
33. **Azziz R** 2003 The evaluation and management of hirsutism. Obstet Gynecol 101:995–1007
34. **Lucky AW, McGuire J, Rosenfield RL, Lucky PA, Rich BH** 1983 Plasma androgens in women with acne vulgaris. J Invest Dermatol 81:70–74
35. **Vexiau P, Husson C, Chivot M, Brerault JL, Fiet J, Julien R, Villette JM, Hardy N, Cathelineau G** 1990 Androgen excess in women with acne alone compared with women with acne and/or hirsutism. J Invest Dermatol 94: 279–283
36. **Slayden SM, Moran C, Sams Jr WM, Boots LR, Azziz R** 2001 Hyperandrogenemia in patients presenting with acne. Fertil Steril 75:889–992
37. **Redmond GP, Olson WH, Lippman JS, Kafrissen ME, Jones TM, Jorizzo JL** 1997 Norgestimate and ethinyl estradiol in the treatment of acne vulgaris: a randomized, placebo-controlled trial. Obstet Gynecol 89:615–622
38. **Thiboutot D, Archer DF, Lemay A, Washenik K, Roberts J, Harrison DD** 2001 A randomized, controlled trial of a low-dose contraceptive containing 20 microg of ethinyl estradiol and 100 microg of levonorgestrel for acne treatment. Fertil Steril 76:461–468
39. **Leyden J, Shalita A, Hordinsky M, Swinyer L, Stanczyk FZ, Weber ME** 2002 Efficacy of a low-dose oral contraceptive containing 20 microg of ethinyl estradiol and 100 microg of levonorgestrel for the treatment of moderate acne: A randomized, placebo-controlled trial. J Am Acad Dermatol 47:399–409
40. **Rosen MP, Breitkopf DM, Nagamani M** 2003 A randomized controlled trial of second- versus third-generation oral contraceptives in the treatment of acne vulgaris. Am J Obstet Gynecol 188:1158–1160
41. **Futterweit W, Dunaif A, Yeh HC, Kingsley P** 1988 The prevalence of hyperandrogenism in 109 consecutive women presenting with alopecia. J Am Acad Dermatol 19:831–836
42. **Carmina E, Lobo RA** 2003 Treatment of hyperandrogenic alopecia in women. Fertil Steril 79:91–95
43. **Carmina E** 1998 Prevalence of idiopathic hirsutism. Eur J Endocrinol 139: 421–423
44. **Balen AH, Conway GS, Kaltsas G, Techatrasak K, Manning PJ, West C, Jacob HS** 1995 Polycystic ovary syndrome: the spectrum of the disorder in 1741 patients. Hum Reprod 10:2107–2111
45. **Pugeat M, Nicolas MH, Craves JC, Alvarado-Dubost C, Fimbel S, Cechaud H, Lejeune H** 1993 Androgens in polycystic ovarian syndrome. Ann NY Acad Sci 687:124–135
46. **Rosner W** 1997 Errors in the measurement of plasma free testosterone. J Clin Endocrinol Metab 82:2014–2015
47. **Vermeulen A, Verdonck L, Kaufman JM** 1999 A critical evaluation of simple methods for the estimation of free testosterone in serum. J Clin Endocrinol Metab 84:3666–3672
48. **Cibula D, Hill M, Starka L** 2000 The best correlation of the new index of hyperandrogenism with the grade of increased hair. Eur J Endocrinol 143: 405–408
49. **Pearlman WH, Crepy O, Murphy M** 1967 Testosterone-binding levels in the serum of women during the normal menstrual cycle, pregnancy, and the postpartum. J Clin Endocrinol Metab 27:1012–1018
50. **Tremblay RR, Dube JY** 1974 Plasma concentration of free and non-TeBG bound testosterone in women on oral contraceptives. Contraception 10: 599–605
51. **Moran C, Knochenhauer ES, Boots LR, Azziz R** 1999 Adrenal androgen excess in hyperandrogenism: relation to age and body mass. Fertil Steril 71:671–674
52. **Meldrum DR, Abraham GE** 1979 Peripheral and ovarian venous concentrations of various steroid hormones in virilizing ovarian tumors. Obstet Gynecol 53:36–43
53. **Friedman CI, Schmidt GE, Kim MH, Powell J** 1985 Serum testosterone concentrations in the evaluation of androgen-producing tumors. Am J Obstet Gynecol 153:44–49
54. **Surrey ES, de Ziegler D, Gambone JC, Judd HL** 1988 Preoperative localization of androgen-secreting tumors: clinical, endocrinologic, and radiologic evaluation of 10 patients. Am J Obstet Gynecol 158:1313–1322
55. **Derksen J, Nagesser SK, Meinders AE, Haak HR, van de Velde CJ** 1994 Identification of virilizing adrenal tumors in hirsute women. N Engl J Med 331:968–973
56. **Ferriman D, Purdie AW** 1983 The aetiology of oligomenorrhoea and/or hirsuties: a study of 467 patients. Postgrad Med J 59:17–20
57. **Christensen SB, Ericsson UB, Janzon L, Tibblin S, Trell E** 1984 The prevalence of thyroid disorders in a middle-aged female population, with special reference to the solitary thyroid nodule. Acta Chir Scand 150:13–19
58. **Schaaf L, Pohl T, Schmidt R, Vardali I, Teuber J, Schlote-Sauter B, Nowotny B, Schiebeler H, Zober A, Usadel KH** 1993 Screening for thyroid disorders in a working population. Clin Investig 71:126–131
59. **Okamura K, Nakashima T, Ueda K, Inoue K, Omae T, Fujishima M** 1987 Thyroid disorders in the general population of Hisayama Japan, with special reference to prevalence and sex differences. Int J Epidemiol 16:545–549
60. **Rallison ML, Dobyns BM, Meikle AW, Bishop M, Lyon JL, Stevens W** 1991 Natural history of thyroid abnormalities: prevalence, incidence, and regression of thyroid diseases in adolescents and young adults. Am J Med 91:363–370
61. **Bjoro T, Holmen J, Kruger O, Midthjell K, Hunstad K, Schreiner T, Sandnes L, Brochmann H** 2000 Prevalence of thyroid disease, thyroid dysfunction and thyroid peroxidase antibodies in a large, unselected population. The Health Study of Nord-Trondelag (HUNT). Eur J Endocrinol 143:639–647
62. **Barbieri RL** 1990 Hyperandrogenic disorders. Clin Obstet Gynecol 33:640–654
63. **Moran C, Tapia M del C, Hernandez E, Vazquez G, Garcia-Hernandez E, Bermudez JA** 1994 Etiological review of hirsutism in 250 patients. Arch Med Res 25:311–314
64. **Zargar AH, Wani AI, Masoodi SR, Laway BA, Bashir MI, Salahuddin M** 2002 Epidemiologic and etiologic aspects of hirsutism in Kashmiri women in the Indian subcontinent. Fertil Steril 77:674–678

JCEM is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.

EXHIBIT S

Cooper, The Cell: A Molecular Approach, Chapter 15 (2000)
available online at
<<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books>>



THE CELL A MOLECULAR APPROACH
SECOND EDITION
GEOFFREY M. COOPER
SINAUER

Short Contents | Full Contents Other books @ NCBI

Navigation

About this book

IV. Cell Regulation

15. Cancer

➡ The Development and Causes of Cancer

Tumor Viruses

Oncogenes

Tumor Suppressor Genes

Applications of Molecular Biology to Cancer Prevention and Treatment

Summary

Questions

References and Further Reading

Figures

Figure 15.1. A malignant tumor of the...

Figure 15.2. Tumor clonality

Figure 15.3. Increased rate of colon cancer...

Figure 15.4. Stages of tumor development

Figure 15.5. Development of colon carcinomas

The Cell ➡ IV. Cell Regulation ➡ 15. Cancer

The Development and Causes of Cancer

The fundamental abnormality resulting in the development of cancer is the continual unregulated proliferation of cancer cells. Rather than responding appropriately to the signals that control normal cell behavior, cancer cells grow and divide in an uncontrolled manner, invading normal tissues and organs and eventually spreading throughout the body. The generalized loss of growth control exhibited by cancer cells is the net result of accumulated abnormalities in multiple cell regulatory systems and is reflected in several aspects of cell behavior that distinguish cancer cells from their normal counterparts.

Types of Cancer

Cancer can result from abnormal proliferation of any of the different kinds of cells in the body, so there are more than a hundred distinct types of cancer, which can vary substantially in their behavior and response to treatment. The most important issue in cancer pathology is the distinction between benign and malignant tumors (Figure 15.1). A tumor is any abnormal proliferation of cells, which may be either benign or malignant. A benign tumor, such as a common skin wart, remains confined to its original location, neither invading surrounding normal tissue nor spreading to distant body sites. A malignant tumor, however, is capable of both invading surrounding normal tissue and spreading throughout the body via the circulatory or lymphatic systems (metastasis). Only malignant tumors are properly referred to as cancers, and it is their ability to invade and metastasize that makes cancer so dangerous. Whereas benign tumors can usually be removed surgically, the spread of malignant tumors to distant body sites frequently makes them resistant to such localized treatment.

Both benign and malignant tumors are classified according to the type of cell from which they arise. Most cancers fall into one of three main groups: carcinomas, sarcomas, and leukemias or lymphomas. Carcinomas, which include approximately 90% of human cancers, are malignancies of epithelial cells. Sarcomas, which are rare in humans, are solid tumors of connective tissues, such as muscle, bone, cartilage, and fibrous tissue. Leukemias and lymphomas, which account for approximately 8% of human malignancies, arise from the blood-forming cells and from cells of the immune system, respectively. Tumors are further classified according to tissue of origin (e.g., lung or breast carcinomas) and the type of cell involved. For example,

Figure 15.6. Structure of representative chemical carcinogens
Figure 15.7. Induction of tumors in mouse...
Figure 15.8. Density-dependent inhibition
Figure 15.9. Autocrine growth stimulation
Figure 15.10. Contact inhibition
Figure 15.11. Defective differentiation and leukemia
Figure 15.12. The focus assay

Tables

Table 15.1. Ten Most Frequent Cancers in...

Search
Go
This book All books
PubMed

fibrosarcomas arise from fibroblasts, and erythroid leukemias from precursors of erythrocytes (red blood cells).

Although there are many kinds of cancer, only a few occur frequently (Table 15.1). More than a million cases of cancer are diagnosed annually in the United States, and more than 500,000 Americans die of cancer each year. Cancers of 10 different body sites account for more than 75% of this total cancer incidence. The four most common cancers, accounting for more than half of all cancer cases, are those of the breast, prostate, lung, and colon/rectum. Lung cancer, by far the most lethal, is responsible for nearly 30% of all cancer deaths.

The Development of Cancer

One of the fundamental features of cancer is tumor clonality, the development of tumors from single cells that begin to proliferate abnormally. The single-cell origin of many tumors has been demonstrated by analysis of X chromosome inactivation (Figure 15.2). As discussed in Chapter 8, one member of the X chromosome pair is inactivated by being converted to heterochromatin in female cells. X inactivation occurs randomly during embryonic development, so one X chromosome is inactivated in some cells, while the other X chromosome is inactivated in other cells. Thus, if a female is heterozygous for an X chromosome gene, different alleles will be expressed in different cells. Normal tissues are composed of mixtures of cells with different inactive X chromosomes, so expression of both alleles is detected in normal tissues of heterozygous females. In contrast, tumor tissues generally express only one allele of a heterozygous X chromosome gene. The implication is that all of the cells constituting such a tumor were derived from a single cell of origin, in which the pattern of X inactivation was fixed before the tumor began to develop.

The clonal origin of tumors does not, however, imply that the original progenitor cell that gives rise to a tumor has initially acquired all of the characteristics of a cancer cell. On the contrary, the development of cancer is a multistep process in which cells gradually become malignant through a progressive series of alterations. One indication of the multistep development of cancer is that most cancers develop late in life. The incidence of colon cancer, for example, increases more than tenfold between the ages of 30 and 50, and another tenfold between 50 and 70 (Figure 15.3). Such a dramatic increase of cancer incidence with age suggests that most cancers develop as a consequence of multiple abnormalities, which accumulate over periods of many years.

At the cellular level, the development of cancer is viewed as a multistep process involving mutation and selection for cells with progressively increasing capacity for proliferation, survival, invasion, and metastasis (Figure 15.4). The first step in the process, **tumor initiation**, is thought to be the result of a genetic alteration leading to abnormal proliferation of a single cell. Cell proliferation then leads to the outgrowth of a population of clonally derived tumor cells. **Tumor progression** continues as additional mutations

occur within cells of the tumor population. Some of these mutations confer a selective advantage to the cell, such as more rapid growth, and the descendants of a cell bearing such a mutation will consequently become dominant within the tumor population. The process is called clonal selection, since a new clone of tumor cells has evolved on the basis of its increased growth rate or other properties (such as survival, invasion, or metastasis) that confer a selective advantage. Clonal selection continues throughout tumor development, so tumors continuously become more rapid-growing and increasingly malignant.

Studies of colon carcinomas have provided a clear example of tumor progression during the development of a common human malignancy (Figure 15.5). The earliest stage in tumor development is increased proliferation of colon epithelial cells. One of the cells within this proliferative cell population is then thought to give rise to a small benign neoplasm (an adenoma or polyp). Further rounds of clonal selection lead to the growth of adenomas of increasing size and proliferative potential. Malignant carcinomas then arise from the benign adenomas, indicated by invasion of the tumor cells through the basal lamina into underlying connective tissue. The cancer cells then continue to proliferate and spread through the connective tissues of the colon wall. Eventually the cancer cells penetrate the wall of the colon and invade other abdominal organs, such as the bladder or small intestine. In addition, the cancer cells invade blood and lymphatic vessels, allowing them to metastasize throughout the body. ↑ TOP

Causes of Cancer

Substances that cause cancer, called carcinogens, have been identified both by studies in experimental animals and by epidemiological analysis of cancer frequencies in human populations (e.g., the high incidence of lung cancer among cigarette smokers). Since the development of malignancy is a complex multistep process, many factors may affect the likelihood that cancer will develop, and it is overly simplistic to speak of single causes of most cancers. Nonetheless, many agents, including radiation, chemicals, and viruses, have been found to induce cancer in both experimental animals and humans.

Radiation and many chemical carcinogens (Figure 15.6) act by damaging DNA and inducing mutations. These carcinogens are generally referred to as initiating agents, since the induction of mutations in key target genes is thought to be the initial event leading to cancer development. Some of the initiating agents that contribute to human cancers include solar ultraviolet radiation (the major cause of skin cancer), carcinogenic chemicals in tobacco smoke, and aflatoxin (a potent liver carcinogen produced by some molds that contaminate improperly stored supplies of peanuts and other grains). The carcinogens in tobacco smoke (including benzo(*a*)pyrene, dimethylnitrosamine, and nickel compounds) are the major identified causes of human cancer. Smoking is the undisputed cause of 80 to 90% of lung cancers, as well as being implicated in cancers of the oral cavity, pharynx, larynx, esophagus, and other sites. In total, it is estimated that smoking is

responsible for nearly one-third of all cancer deaths—an impressive toll for a single carcinogenic agent.

Other carcinogens contribute to cancer development by stimulating cell proliferation, rather than by inducing mutations. Such compounds are referred to as tumor promoters, since the increased cell division they induce is required for the outgrowth of a proliferative cell population during early stages of tumor development. The phorbol esters that stimulate cell proliferation by activating protein kinase C (see Figure 13.26) are classic examples. Their activity was defined by studies of chemical induction of skin tumors in mice (Figure 15.7). Tumorigenesis in this system can be initiated by a single treatment with a mutagenic carcinogen. Tumors do not develop, however, unless the mice are subsequently treated with a tumor promoter (usually a phorbol ester) to stimulate proliferation of the mutated cells.

Hormones, particularly estrogens, are important as tumor promoters in the development of some human cancers. The proliferation of cells of the uterine endometrium, for example, is stimulated by estrogen, and exposure to excess estrogen significantly increases the likelihood that a woman will develop endometrial cancer. The risk of endometrial cancer is therefore substantially increased by long-term postmenopausal estrogen replacement therapy with high doses of estrogen alone. Fortunately, this risk is minimized by administration of progesterone to counteract the stimulatory effect of estrogen on endometrial cell proliferation. However, long-term therapy with combinations of estrogen and progesterone may lead to an increased risk of breast cancer.

In addition to chemicals and radiation, some viruses induce cancer both in experimental animals and in humans. The common human cancers caused by viruses include liver cancer and cervical carcinoma, which together account for 10 to 20% of worldwide cancer incidence. These viruses are important not only as causes of human cancer; as discussed later in this chapter, studies of tumor viruses have played a key role in elucidating the molecular events responsible for the development of cancers induced by both viral and nonviral carcinogens. ↑ TOP

Properties of Cancer Cells

The uncontrolled growth of cancer cells results from accumulated abnormalities affecting many of the cell regulatory mechanisms that have been discussed in preceding chapters. This relationship is reflected in several aspects of cell behavior that distinguish cancer cells from their normal counterparts. Cancer cells typically display abnormalities in the mechanisms that regulate normal cell proliferation, differentiation, and survival. Taken together, these characteristic properties of cancer cells provide a description of malignancy at the cellular level.

The uncontrolled proliferation of cancer cells *in vivo* is mimicked by their behavior in cell culture. A primary distinction between cancer cells and normal cells in culture is that normal cells display **density-dependent**

inhibition of cell proliferation (Figure 15.8). Normal cells proliferate until they reach a finite cell density, which is determined in part by the availability of growth factors added to the culture medium (usually in the form of serum). They then cease proliferating and become quiescent, arrested in the G_0 stage of the cell cycle (see Figure 14.6). The proliferation of most cancer cells, however, is not sensitive to density-dependent inhibition. Rather than responding to the signals that cause normal cells to cease proliferation and enter G_0, tumor cells generally continue growing to high cell densities in culture, mimicking their uncontrolled proliferation *in vivo*.

A related difference between normal cells and cancer cells is that many cancer cells have reduced requirements for extracellular growth factors. As discussed in Chapter 13, the proliferation of most cells is controlled, at least in part, by polypeptide growth factors. For some cell types, particularly fibroblasts, the availability of serum growth factors is the principal determinant of their proliferative capacity in culture. The growth factor requirements of these cells are closely related to the phenomenon of density-dependent inhibition, since the density at which normal fibroblasts become quiescent is proportional to the concentration of serum growth factors in the culture medium.

The growth factor requirements of many tumor cells are reduced compared to their normal counterparts, contributing to the unregulated proliferation of tumor cells both *in vitro* and *in vivo*. In some cases, cancer cells produce growth factors that stimulate their own proliferation (Figure 15.9). Such abnormal production of a growth factor by a responsive cell leads to continuous autostimulation of cell division (**autocrine growth stimulation**), and the cancer cells are therefore less dependent on growth factors from other, physiologically normal sources. In other cases, the reduced growth factor dependence of cancer cells results from abnormalities in intracellular signaling systems, such as unregulated activity of growth factor receptors or other proteins (e.g., Ras proteins or protein kinases) that were discussed in Chapter 13 as elements of signal transduction pathways leading to cell proliferation.

Cancer cells are also less stringently regulated than normal cells by cell-cell and cell-matrix interactions. Most cancer cells are less adhesive than normal cells, often as a result of reduced expression of cell surface adhesion molecules. For example, loss of E-cadherin, the principal adhesion molecule of epithelial cells, is important in the development of carcinomas (epithelial cancers). As a result of reduced expression of cell adhesion molecules, cancer cells are comparatively unrestrained by interactions with other cells and tissue components, contributing to the ability of malignant cells to invade and metastasize. The reduced adhesiveness of cancer cells also results in morphological and cytoskeletal alterations: Many tumor cells are rounder than normal, in part because they are less firmly attached to either the extracellular matrix or neighboring cells.

A striking difference in the cell-cell interactions displayed by normal cells and those of cancer cells is illustrated by the phenomenon of contact

inhibition (Figure 15.10). Normal fibroblasts migrate across the surface of a culture dish until they make contact with a neighboring cell. Further cell migration is then inhibited, and normal cells adhere to each other, forming an orderly array of cells on the culture dish surface. Tumor cells, in contrast, continue moving after contact with their neighbors, migrating over adjacent cells, and growing in disordered, multilayered patterns. Not only the movement but also the proliferation of many normal cells is inhibited by cell-cell contact, and cancer cells are characteristically insensitive to such contact inhibition of growth.

Two additional properties of cancer cells affect their interactions with other tissue components, thereby playing important roles in invasion and metastasis. First, malignant cells generally secrete proteases that digest extracellular matrix components, allowing the cancer cells to invade adjacent normal tissues. Secretion of collagenase, for example, appears to be an important determinant of the ability of carcinomas to digest and penetrate through basal laminae to invade underlying connective tissue (see Figure 15.5). Second, cancer cells secrete growth factors that promote the formation of new blood vessels (angiogenesis). Angiogenesis is needed to support the growth of a tumor beyond the size of about a million cells, at which point new blood vessels are required to supply oxygen and nutrients to the proliferating tumor cells. Such blood vessels are formed in response to growth factors, secreted by the tumor cells, that stimulate proliferation of endothelial cells in the walls of capillaries in surrounding tissue, resulting in the outgrowth of new capillaries into the tumor. The formation of such new blood vessels is important not only in supporting tumor growth, but also in metastasis. The actively growing new capillaries formed in response to angiogenic stimulation are easily penetrated by the tumor cells, providing a ready opportunity for cancer cells to enter the circulatory system and begin the metastatic process.

Another general characteristic of most cancer cells is that they fail to differentiate normally. Such defective differentiation is closely coupled to abnormal proliferation, since, as discussed in Chapter 14, most fully differentiated cells either cease cell division or divide only rarely. Rather than carrying out their normal differentiation program, cancer cells are usually blocked at an early stage of differentiation, consistent with their continued active proliferation.

The leukemias provide a particularly good example of the relationship between defective differentiation and malignancy. All of the different types of blood cells are derived from a common stem cell in the bone marrow (see Figure 14.44). Descendants of these cells then become committed to specific differentiation pathways. Some cells, for example, differentiate to form erythrocytes whereas others differentiate to form lymphocytes, granulocytes, or macrophages. Cells of each of these types undergo several rounds of division as they differentiate, but once they become fully differentiated, cell division ceases. Leukemic cells, in contrast, fail to undergo terminal differentiation (Figure 15.11). Instead, they become arrested at early stages of maturation at which they retain their capacity for proliferation and

continue to reproduce.

As discussed in Chapter 13, programmed cell death, or apoptosis, is an integral part of the differentiation program of many cell types, including blood cells. Many cancer cells fail to undergo apoptosis, and therefore exhibit increased life spans compared to their normal counterparts. This failure of cancer cells to undergo programmed cell death contributes substantially to tumor development. For example, the survival of many normal cells is dependent on signals from growth factors or from the extracellular matrix that prevent apoptosis. In contrast, tumor cells are often able to survive in the absence of growth factors required by their normal counterparts. Such a failure of tumor cells to undergo apoptosis when deprived of normal environmental signals may be important not only in primary tumor development but also in the survival and growth of metastatic cells in abnormal tissue sites. Normal cells also undergo apoptosis following DNA damage, while many cancer cells fail to do so. In this case, the failure to undergo apoptosis contributes to the resistance of cancer cells to irradiation and many chemotherapeutic drugs, which act by damaging DNA. Abnormal cell survival, as well as cell proliferation, thus plays a major role in the unrelenting growth of cancer cells in an animal. TOP

Transformation of Cells in Culture

The study of tumor induction by radiation, chemicals, or viruses requires experimental systems in which the effects of a carcinogenic agent can be reproducibly observed and quantitated. Although the activity of carcinogens can be assayed in intact animals, such experiments are difficult to quantitate and control. The development of *in vitro* assays to detect the conversion of normal cells to tumor cells in culture, a process called **cell transformation**, therefore represented a major advance in cancer research. Such assays are designed to detect transformed cells, which display the *in vitro* growth properties of tumor cells, following exposure of a culture of normal cells to a carcinogenic agent. Their application has allowed experimental analysis of cell transformation to reach a level of sophistication that could not have been attained by studies in whole animals alone.

The first and most widely used assay of cell transformation is the focus assay, which was developed by Howard Temin and Harry Rubin in 1958. The focus assay is based on the ability to recognize a group of transformed cells as a morphologically distinct "focus" against a background of normal cells on the surface of a culture dish (Figure 15.12). The focus assay takes advantage of three properties of transformed cells: altered morphology, loss of contact inhibition, and loss of density-dependent inhibition of growth. The result is the formation of a colony of morphologically altered transformed cells that overgrow the background of normal cells in the culture. Such foci of transformed cells can usually be detected and quantified within a week or two after exposure to a carcinogenic agent. In general, cells transformed *in vitro* are able to form tumors following inoculation into susceptible animals, supporting *in vitro* transformation as a valid indicator of the formation of cancer cells. TOP

EXHIBIT T

Wang et al, 2004, Science 304:1164

ancestor of domestic dogs. Thus, dogs from these breeds may be the best living representatives of the ancestral dog gene pool. It is notable that several breeds commonly believed to be of ancient origin, such as the Pharaoh Hound and Ibizan Hound, are not included in this group. These are often thought to be the oldest of all dog breeds, descending directly from the ancient Egyptian dogs drawn on tomb walls more than 5000 years ago. Our results indicate, however, that these two breeds have been recreated in more recent times from combinations of other breeds. Thus, although their appearance matches the ancient Egyptian sight hounds, their genomes do not. Similar conclusions apply to the Norwegian Elkhound, which clusters with modern European breeds rather than with the other Arctic dogs, despite reports of direct descent from Scandinavian origins more than 5000 years ago (*1*, *24*).

The large majority of breeds appears to represent a more recent radiation from shared European stock. Although the individual breeds are genetically differentiated, they appear to have diverged at essentially the same time. This radiation probably reflects the proliferation of distinct breeds from less codified phenotypic varieties after the introduction of the breed concept and the creation of breed clubs in Europe in the 1800s. A more sensitive cluster analysis was able to discern additional genetic structure of three subpopulations within this group. One contains Mastiff-like breeds and appears to reflect shared morphology derived from a common ancestor. Another includes Shetland Sheepdog, the two Belgian Sheepdogs, and Collie, and may reflect shared ancestral herding behavior. The remaining population is dominated by a proliferation of breeds dedicated to various aspects of the hunt. For these breeds, historical and breed club records suggest highly intertwined bloodlines, consistent with our results.

Dog breeds have traditionally been grouped on the basis of their roles in human activities, physical phenotypes, and historical records. Here, we defined an independent classification based on patterns of genetic variation. This classification supports a subset of traditional groupings and also reveals previously unrecognized connections among breeds. An accurate understanding of the genetic relationships among breeds lays the foundation for studies aimed at uncovering the complex genetic basis of breed differences in morphology, behavior, and disease susceptibility.

References and Notes

1. J. Crowley, B. Adelman, Eds., *The Complete Dog Book; Official Publication of the American Kennel Club* (Howell Book House, New York, ed. 19, 1998).
2. D. F. Patterson *et al.*, *J. Am. Vet. Med. Assoc.* **193**, 1131 (1988).
3. E. A. Ostrander, L. Kruglyak, *Genome Res.* **10**, 1271 (2000).
4. C. Vila *et al.*, *Science* **276**, 1687 (1997).
5. P. Savolainen, Y. P. Zhang, J. Luo, J. Lundeberg, T. Leitner, *Science* **298**, 1610 (2002).
6. J. A. Leonard *et al.*, *Science* **298**, 1613 (2002).
7. C. A. Rogers, A. H. Brace, *The International Encyclopedia of Dogs* (Howell Book House, New York, ed. 1, 1995).
8. B. Fogel, *The Encyclopedia of the Dog* (DK Publishing, New York, 1995).
9. M. T. Koskinen, *Anim. Genet.* **34**, 297 (2003).
10. D. N. Irion *et al.*, *J. Hered.* **94**, 81 (2003).
11. J. K. Pritchard, M. Stephens, P. Donnelly, *Genetics* **155**, 945 (2000).
12. D. Falush, M. Stephens, J. K. Pritchard, *Genetics* **164**, 1567 (2003).
13. F. Tajima, M. Nei, *Mol. Biol. Evol.* **1**, 269 (1984).
14. R. Sachidanandam *et al.*, *Nature* **409**, 928 (2001).
15. J. C. Venter *et al.*, *Science* **291**, 1304 (2001).
16. L. Excoffier, P. E. Smouse, J. M. Quattro, *Genetics* **131**, 479 (1992).
17. N. A. Rosenberg *et al.*, *Science* **298**, 2381 (2002).
18. L. L. Cavelli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1994).
19. D. E. MacHugh, R. T. Loftus, P. Cunningham, D. G. Bradley, *Anim. Genet.* **29**, 333 (1998).
20. G. Laval *et al.*, *Genet. Sel. Evol.* **32**, 187 (2000).
21. J. K. Pritchard, M. Stephens, N. A. Rosenberg, P. Donnelly, *Am. J. Hum. Genet.* **67**, 170 (2000).
22. Materials and methods are available as supporting material on *Science* Online.
23. T. Yamazaki, K. Yamazaki, *Legacy of the Dog: The Ultimate Illustrated Guide to Over 200 Breeds* (Chronicle Books, San Francisco, CA, 1995).
24. B. Wilcox, C. Walkowicz, *Atlas of Dog Breeds of the World* (TFH Publications, Neptune City, NJ, ed. 5, 1995).
25. D. Paetkau, W. Calvert, I. Stirling, C. Strobeck, *Mol. Ecol.* **4**, 347 (1995).
26. L. L. Cavalli-Sforza, A. W. Edwards, *Evolution* **32**, 550 (1967).
27. D. B. Goldstein, A. R. Linares, L. L. Cavalli-Sforza, M. W. Feldman, *Genetics* **139**, 463 (1995).
28. N. A. Rosenberg *et al.*, *Genetics* **159**, 699 (2001).
29. D. Falush *et al.*, *Science* **299**, 1582 (2003).
30. M. V. Sablin, G. A. Khlopachev, *Curr. Anthropol.* **43**, 795 (2002).

31. Supported by the Burroughs Wellcome Innovation Award (E.A.O. and L.K.), grants from the AKC–Canine Health Foundation (G.S.J.), NIH training grant T32 HG00035 (H.G.P.), and a postdoctoral fellowship from the Waltham Foundation (N.B.S.). E.A.O. also acknowledges support from K05 CA90754. L.K. is a James S. McDonnell Centennial Fellow. We thank the many dog owners, researchers, and breeders who provided DNA samples for this work, especially G. Brewer, C. Gaiser, and K. Murphy; D. Lynch, A. Ziska, C. Ramirez, M. Langlois, and D. Akey for assistance with sample collection; M. Stephens for advice and assistance with the program *structure*; K. Markianos and J. Akey for helpful discussions; M. Eberle for computing assistance; H. Coller, E. Giniger, R. Wayne, and three anonymous reviewers for comments on the manuscript; and the AKC and C. Jierski for use of the canine artwork included in Fig. 3.

Supporting Online Material

www.sciencemag.org/cgi/content/full/304/5674/1160/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S5
References

2 March 2004; accepted 21 April 2004

Mutational Analysis of the Tyrosine Phosphatome in Colorectal Cancers

Zhenghe Wang,[1] Dong Shen,[1]* D. Williams Parsons,[1]* Alberto Bardelli,[1]* Jason Sager,[1] Steve Szabo,[1] Janine Ptak,[1] Natalie Silliman[1], Brock A. Peters,[1] Michiel S. van der Heijden,[1] Giovanni Parmigiani,[1] Hai Yan,[2] Tian-Li Wang,[1] Greg Riggins,[1] Steven M. Powell,[3] James K. V. Willson,[4] Sanford Markowitz,[4] Kenneth W. Kinzler,[1] Bert Vogelstein,[1] Victor E. Velculescu[1]†

Tyrosine phosphorylation, regulated by protein tyrosine phosphatases (PTPs) and kinases (PTKs), is important in signaling pathways underlying tumorigenesis. A mutational analysis of the tyrosine phosphatase gene superfamily in human cancers identified 83 somatic mutations in six PTPs (*PTPRF, PTPRG, PTPRT, PTPN3, PTPN13, PTPN14*), affecting 26% of colorectal cancers and a smaller fraction of lung, breast, and gastric cancers. Fifteen mutations were nonsense, frameshift, or splice-site alterations predicted to result in truncated proteins lacking phosphatase activity. Five missense mutations in the most commonly altered PTP (*PTPRT*) were biochemically examined and found to reduce phosphatase activity. Expression of wild-type but not a mutant PTPRT in human cancer cells inhibited cell growth. These observations suggest that the mutated tyrosine phosphatases are tumor suppressor genes, regulating cellular pathways that may be amenable to therapeutic intervention.

Phosphorylation of tyrosine residues is a central feature of many cellular signaling pathways, including those affecting growth, differentiation, cell cycle regulation, apoptosis, and invasion (*1*, *2*). This phosphorylation is coordinately controlled by protein tyrosine kinases (PTKs) and phosphatases (PTPs). Although a variety of PTK genes have been directly linked to tumorigenesis through somatic activating mutations (*3–6*), only a few PTP genes have been implicated in cancer (*7–10*). Moreover, it is not known how many or how frequently members

of the PTP gene family are altered in any particular cancer type. We have systematically addressed these issues by comprehensive mutational analysis of the PTP gene superfamily in colorectal tumors.

The PTP gene superfamily is composed of three main families: (i) the classical PTPs, including the receptor PTPs (RPTPs) and the nonreceptor PTPs (NRPTPs); (ii) the dual specificity phosphatases (DSPs), which can dephosphorylate serine and threonine in addition to tyrosine residues; and (iii) the low molecular weight phosphatases (LMPs) (*1*). Using an approach similar to that described in a recent bioinformatic analysis of PTPs in the mouse genome (*11*), we employed a combination of Hidden Markov Models representing catalytic domains of members of the PTP superfamily to identify 53 classical PTPs (21 RPTPs and 32 NRPTPs), 33 DSPs, and one LMP in the human genome (*12*). This analysis revealed a set of genes representing all known human PTPs (*13*) as well as seven putative PTPs.

As an initial screen to evaluate whether these phosphatases are genetically altered in human cancer, we analyzed the coding exons of all 87 members of this gene superfamily in 18 colorectal cancers. A total of 1375 exons from all annotated RPTPs, NRPTPs, DSPs, and LMPs were extracted from genomic databases (*12*). These exons were amplified by polymerase chain reaction from cancer genomic DNA samples and directly sequenced with dye terminator chemistry (*12*). Whenever a presumptive mutation was identified, we attempted to determine whether it was somatically acquired (i.e., tumor specific) by examining the sequence of the gene in genomic DNA from normal tissue of the relevant patient.

From the 3.3 Mb of sequence information obtained, we identified six genes containing somatic mutations, including three members of the RPTP subfamily (*PTPRF*, *PTPRG*, and *PTPRT*) and three members of the NRPTP subfamily (*PTPN3*, *PTPN13*, and *PTPN14*). These six genes were then further analyzed for mutations in another 157 colorectal cancers. Through this strategy, we identified 77 mutations in the six genes, in aggregate affecting 26% of the colorectal tumors analyzed (table S1 and Fig. 1). Examination of these six genes in seven other tumor types identified *PTPRT* mutations in two of 11 (18%) lung cancers and two of 12 gastric cancers (17%), and *PTPRF* mutations in one of 11 (9%) lung cancers and one of 11 (9%) breast cancers. No mutations were identified in 12 pancreatic cancers, 12 ovarian cancers, 12 medulloblastomas, or 12 glioblastomas (table S1 and Fig. 1). In total, 83 nonsynonymous mutations were observed, all of which were somatic in the cancers that could be assessed (*12*).

Fifteen of the 83 mutations were nonsense, frameshift, or splice-site alterations, all of which were predicted to result in aberrant or truncated proteins. In 16 tumors both alleles of the phosphatase gene appeared to be mutated, a characteristic often associated with tumor suppressor genes. The majority of tumors with PTP gene mutations also contained mutations in *KRAS* or *BRAF*, and nine tumors contained alterations in previously reported tyrosine kinase genes (table S1). Thus, the mutant phosphatases identified in this study are likely to operate through cellular pathways distinct from those associated with previously identified mutant kinases.

Analysis of mutations in tumors is complicated by the fact that mutations can arise either as functional alterations affecting key genes underlying the neoplastic process or as nonfunctional "passenger" changes. The multiple waves of clonal expansion and selection that occur throughout tumorigenesis lead to fixation of any mutation that had previously occurred in any predecessor cell, regardless of whether the mutation was actually responsible for the clonal expansion. Two independent lines of evidence suggest that the sequence alterations we observed are functional. First, the ratio of nonsynonymous to synonymous mutations provides an indication of selection, as synonymous alterations usually do not exert a growth advantage. There were no somatic synonymous mutations detected in the colorectal cancers analyzed, resulting in a ratio of nonsynonymous to synonymous mutations of 77 to 0, much higher than the expected 2:1 ratio for nonselected passenger mutations ($P < 1 \times 10^{-6}$). Second, the prevalence of mutations in the coding regions of the analyzed genes was ~19 per Mb of tumor DNA, similar to the prevalence of functional somatic alterations observed in other gene families [e.g., the tyrosine kinome (*6*)] and significantly higher than the prevalence of nonfunctional alterations previously observed in the cancer genomes (~1 per Mb, $P < 0.01$) (*14*). These data support the idea that these mutations were the targets of selection during tumorigenesis.

The great majority of the nonsense and frameshift mutations (Fig. 1) would result in polypeptides devoid of C-terminal phosphatase catalytic domains, thereby leading to inactivation of the phosphatase. To evaluate whether tumor-specific point mutations alter phosphatase activity, we biochemically tested mutant versions of PTPRT, the most frequently mutated PTP in the superfamily. Mutations in both intracellular PTP domains (D1 and D2) were evaluated. His-tagged versions of the catalytic region of wild-type PTPRT, two D1 mutants (Q987K and N1128I), and three D2 mutants (R1212W, R1346L, and T1368M) were produced in bacteria and analyzed for phosphatase activity by using 6,8-difluoro-4-methylumbelliferyl phosphate (DiFMUP) as a substrate (Fig. 2) (*12*). All D1 and D2 mutants had reduced phosphatase activity compared with the wild-type protein (Fig. 2). The kinetic parameter K_{cat} was reduced in both D1 mutants, while K_m was increased in all three D2 mutants, suggesting that mutations in the two domains may have different effects on enzy-

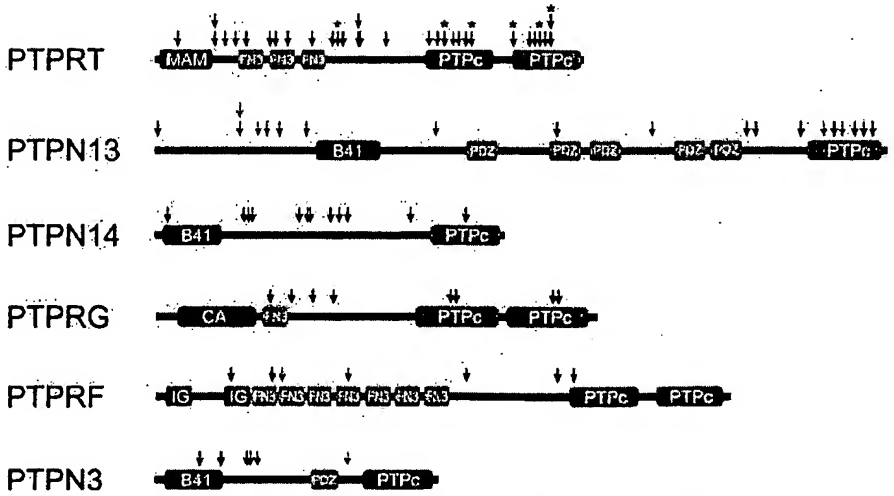


Fig. 1. Distribution of mutations in PTPRT, PTPN13, PTPN14, PTPRG, PTPRF, and PTPN3. Black arrows indicate location of missense mutations, red arrows indicate location of nonsense mutations or frameshifts, and boxes represent functional domains (B41, band 41; CA, carbonic anhydrase; FN3, fibronectin type III; IG, immunoglobulin; MAM, meprin/A5/PTPμ; PDZ, postsynaptic density, discs large, zonula occludans; PTPc, catalytic phosphatase domain). Black stars indicate PTPRT mutants evaluated for phosphatase activity (Fig. 2), and the red star indicates PTPRT mutant evaluated for effects on cell proliferation (Fig. 3).

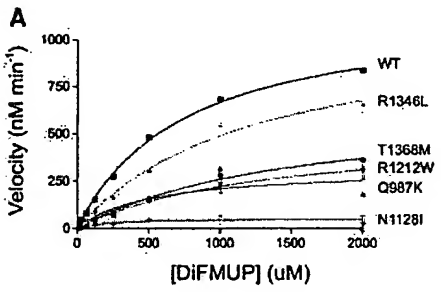
[1]The Sidney Kimmel Comprehensive Cancer Center, The Howard Hughes Medical Institute, The Johns Hopkins University Medical Institutions, Baltimore, MD 21231, USA. [2]Department of Pathology, Duke University Medical Center, Durham, NC 27710, USA. [3]Division of Gastroenterology/Heptalogy, UVA Health System, Charlottesville, VA 22908, USA. [4]Howard Hughes Medical Institute and Department of Medicine and Ireland Cancer Center, University Hospitals of Cleveland and Case Western Reserve University, Cleveland, OH 44106, USA

*These authors contributed equally to this work.
†To whom correspondence should be addressed: E-mail: velculescu@jhmi.edu

matic activity. Although it has been thought that the D2 domain is usually catalytically inactive (*1*), our results are consistent with recent studies that show that the D2 domain is important for phosphatase activity in some receptor phosphatases (*15*).

These biochemical data on missense mutations, coupled with the large number of truncating mutations, suggested that *PTPRT* functions as a tumor suppressor gene. To determine whether PTPRT expression can inhibit tumor cell growth, we transfected wild-type PTPRT into HCT116 colorectal cancer cells (*12*). An identical expression vector containing an R632X mutant of PTPRT was used for comparison. Wild-type PTPRT potently inhibited cell growth in this assay, as seen by the substantial decrease in the number of neomycin-resistant colonies compared with the R632X mutant or with vector alone (Fig. 3, A and B). Similar results with wild-type and mutant PTPRT were also observed in DLD1 colorectal cancer cells.

The combination of these genetic, biochemical, and cellular data suggest that *PTPRT* and the other identified phosphatases are likely to act as tumor suppressors. This is consistent with the function of other phosphatases implicated in tumorigenesis (*7*, *8*, *16*) and with the general role of phosphatases in inhibiting various growth-promoting signaling pathways (*2*).



B

PTPRT protein	K_m (uM)	K_{cat} (min^{-1})	K_{cat}/K_m (M^{-1} min^{-1})
WT	753	11.7	15,512
Q987K mutant	497	3.1	6,313
N1128I mutant	207	0.5	2,570
R1212W mutant	1,289	5.1	3,936
R1346L mutant	1,229	10.9	8,869
T1368M mutant	1,499	6.5	4,317

Fig. 2. Evaluation of phosphatase activity of mutant PTPRT. (A) Saturation kinetics of wild-type and mutant PTPRT. His-tagged versions of PTPRT protein segments comprising the two catalytic domains containing wild-type (WT) and tumor-specific mutant sequences were expressed in bacteria and purified using nickel-affinity chromatography. Equal amounts of WT and mutant proteins were used to evaluate enzyme kinetics. The rate of hydrolysis of substrate (DiFMUP) is plotted against increasing substrate concentration. Data were fitted to the Michaelis-Menton equation, and the resulting kinetic parameters of WT and mutant proteins are indicated in (B).

The absence of biallelic mutations in a subset of the analyzed tumors suggests that some alterations may act in a dominant-negative fashion or may affect gene dosage, mechanisms that have been previously involved in inactivation of other tumor suppressor genes (*17*, *18*).

Little is known about the functional role of the tyrosine phosphatases discussed here. PTPN13 appears to be involved in apoptosis (*19*) and may be partly responsible for the antitumor effects of tamoxifen (*20*). Overexpression of PTPN3 inhibits growth of NIH/3T3 cells, possibly through interaction with valosin-containing protein (VCP/p97) (*21*). PTPN14 and PTPRF are thought to play a role in cell adhesion by regulating tyrosine phosphorylation of adherens junction proteins (*22*, *23*). Because increased phosphorylation of adherens junctions has been shown to increase cell motility and migration (*22*, *24*), mutational inactivation of these genes may be an important step in cancer cell invasion and metastasis. *PTPRG* maps to chromosome 3p14.2, a region frequently lost in lung, renal, and early-stage breast tumors, and is thought to be a target of the translocation at 3p14 in familial renal cell carcinoma (*25*–*27*). However, no point mutations in *PTPRG* (*28*) or any of the other genes identified here have been previously described in any cancer. PTPRT is expressed in the developing central nervous system and in the adult cerebellum (*29*) and had not been thought to play a role in the growth or differentiation of other tissues. We have found that PTPRT is expressed in a variety of human tissues, including normal colon epithelium as well as cells derived from colorectal cancers (fig. S1).

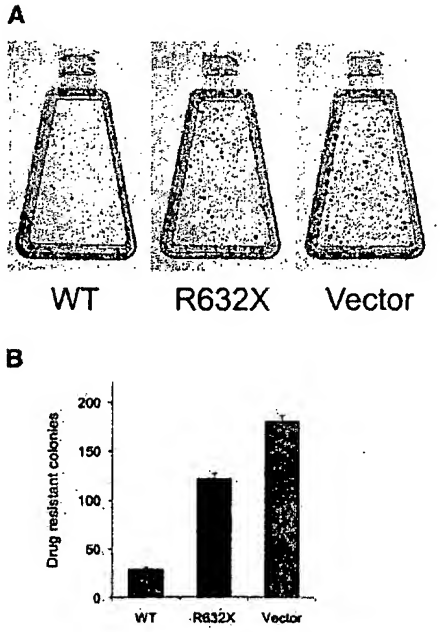


Fig. 3. PTPRT overexpression suppresses growth of human cancer cells. (A) HCT116 colorectal cancer cells were transfected with wild-type (WT) PTPRT construct, truncated R632X mutant PTPRT construct, or empty pCI-Neo vector. The photographs show colonies stained with crystal violet after 14 days of geneticin selection. (B) Number of resistant colonies (mean of two 25 cm^2 flasks) for WT PTPRT, mutant PTPRT, and empty vector.

Phosphatases affect signaling pathways that may be amenable to therapeutic intervention in cancer cells (*2*). Although reactivation of incapacitated phosphatases is likely to be pharmacologically challenging, identification of the corresponding kinases that phosphorylate substrates normally regulated by the mutant phosphatases could provide novel therapeutic targets. Like the analysis of genetic alterations in tyrosine kinases (*6*), the present study suggests the possibility of individualized therapy based on the mutant phosphatases present in specific tumors. Conceivably, this approach would have broad therapeutic use, because more than 50% of colorectal tumors analyzed to date have alterations in at least one member of the tyrosine phosphatome or kinome.

References and Notes

1. B. G. Neel, N. K. Tonks, *Curr. Opin. Cell Biol.* **9**, 193 (1997).
2. T. Hunter, *Philos. Trans. R. Soc. London B Ser.* **353**, 583 (1998).
3. P. Blume-Jensen, T. Hunter, *Nature* **411**, 355 (2001).
4. H. Davies *et al.*, *Nature* **417**, 949 (2002).
5. H. Rajagopalan *et al.*, *Nature* **418**, 934 (2002).
6. A. Bardelli *et al.*, *Science* **300**, 949 (2003).
7. J. Li *et al.*, *Science* **275**, 1943 (1997).
8. P. A. Steck *et al.*, *Nature Genet.* **15**, 356 (1997).
9. M. Tartaglia *et al.*, *Nature Genet.* **34**, 148 (2003).
10. S. Saha *et al.*, *Science* **294**, 1343 (2001).
11. A. R. Forrest *et al.*, *Genome Res.* **13**, 1443 (2003).
12. Materials and methods are available as supporting material on *Science* Online.
13. J. N. Andersen *et al.*, *Mol. Cell. Biol.* **21**, 7117 (2001).
14. T. L. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3076 (2002).
15. A. Oganesian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7563 (2003).
16. S. S. Wang *et al.*, *Science* **282**, 284 (1998).
17. S. E. Kern *et al.*, *Science* **256**, 827 (1992).
18. R. Fodde, R. Smits, *Science* **298**, 761 (2002).
19. G. Bompard, C. Puech, C. Prebois, F. Vignon, G. Freiss, *J. Biol. Chem.* **277**, 47861 (2002).
20. G. Freiss, C. Puech, F. Vignon, *Mol. Endocrinol.* **12**, 568 (1998).
21. S. H. Zhang, J. Liu, R. Kobayashi, N. K. Tonks, *J. Biol. Chem.* **274**, 17806 (1999).
22. C. Wadham, J. R. Gamble, M. A. Vadas, Y. Khew-Goodall, *Mol. Biol. Cell* **14**, 2520 (2003).
23. T. Muller, A. Choidas, E. Reichmann, A. Ullrich, *J. Biol. Chem.* **274**, 10173 (1999).
24. O. Ayalon, B. Geiger, *J. Cell Sci.* **110**, 547 (1997).
25. S. LaForgia *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5036 (1991).
26. I. Panagopoulos *et al.*, *Cancer Res.* **56**, 4871 (1996).
27. K. Kastury *et al.*, *Genomics* **32**, 225 (1996).
28. T. Druck *et al.*, *Cancer Res.* **55**, 5348 (1995).
29. P. E. McAndrew *et al.*, *J. Comp. Neurol.* **391**, 444 (1998).
30. We thank S. Kern and A. Gazdar for providing tumor samples. Supported by The Ludwig Trust, The Maryland Cigarette Restitution Fund, The Benjamin Baker Scholarship Fund, The Clayton Fund, and NIH grants CA 43460, CA 57345, and CA 62924.

Supporting Online Material

www.sciencemag.org/cgi/content/full/304/5674/1164/DC1
Materials and Methods
Fig. S1
Tables S1 to S3
References

27 January 2004; accepted 12 April 2004

EXHIBIT U

Wadham et al., Molecular Biology of the Cell 14:2520 (2003)

Molecular Biology of the Cell
Vol. 14, 2520–2529, June 2003

The Protein Tyrosine Phosphatase Pez Is a Major Phosphatase of Adherens Junctions and Dephosphorylates β-Catenin

Carol Wadham,* Jennifer R Gamble,*† Mathew A Vadas,*† and Yeesim Khew-Goodall*‡

*Hanson Centre for Cancer Research, Institute of Medical and Veterinary Science, Adelaide, SA 5000, Australia; and †The University of Adelaide, Adelaide, SA 5005, Australia

Submitted September 10, 2002; Revised January 13, 2003; Accepted January 30, 2003
Monitoring Editor: Mark H. Ginsberg

Cell-cell adhesion regulates processes important in embryonal development, normal physiology, and cancer progression. It is regulated by various mechanisms including tyrosine phosphorylation. We have previously shown that the protein tyrosine phosphatase Pez is concentrated at intercellular junctions in confluent, quiescent monolayers but is nuclear in cells lacking cell-cell contacts. We show here with an epithelial cell model that Pez localizes to the adherens junctions in confluent monolayers. A truncation mutant lacking the catalytic domain acts as a dominant negative mutant to upregulate tyrosine phosphorylation at adherens junctions. We identified β-catenin, a component of adherens junctions, as a substrate of Pez by a "substrate trapping" approach and by in vitro dephosphorylation with recombinant Pez. Consistent with this, ectopic expression of the dominant negative mutant caused an increase in tyrosine phosphorylation of β-catenin, demonstrating that Pez regulates the level of tyrosine phosphorylation of adherens junction proteins, including β-catenin. Increased tyrosine phosphorylation of adherens junction proteins has been shown to decrease cell-cell adhesion, promoting cell migration as a result. Accordingly, the dominant negative Pez mutant enhanced cell motility in an in vitro "wound" assay. This suggests that Pez is also a regulator of cell motility, most likely through its action on cell-cell adhesion.

INTRODUCTION

Cell-cell adhesion regulates diverse cellular functions including cell proliferation, migration, and apoptosis (Vleminckx and Kemler, 1999). One important cell-cell adhesion system, the adherens junction (AJ), is mediated by a family of homophilic receptors, the cadherins (Steinberg and McNutt, 1999). The strength of cadherin-mediated adhesion is regulated by lateral clustering of cadherin molecules at the plasma membrane and also through the linkage of its intracellular cytoplasmic tail to the actin cytoskeleton. β-Catenin, a structural component of AJs and signal transducer of the wnt signaling pathway, is crucial for cross-linking cadherins to the actin cytoskeleton through another intermediate, α-catenin (Gumbiner, 1995;Cowin and Burke, 1996).

Article published online ahead of print. Mol. Biol. Cell 10.1091/mbc.E02-09-0577. Article and publication date are available at www.molbiolcell.org/cgi/doi/10.1091/mbc.E02-09-0577.

‡Corresponding author. E-mail address: yeesim.khew-goodall@imvs.sa.gov.au.

Reversible tyrosine phosphorylation, catalyzed by the opposing actions of protein tyrosine kinases (PTKs) and protein tyrosine phosphatases (PTPs), is an important mechanism for regulating the linkage of cadherins to the cytoskeleton. A number of PTKs and PTPs have been found to be associated with AJs (Steinberg and McNutt, 1999). Inhibitors of PTPs have been shown to disrupt cell-cell adhesion, suggesting that PTPs play a critical role in maintaining the integrity of AJs (Ayalon and Geiger, 1997). The observation that phosphorylation of a critical tyrosine residue, Tyr654, on β-catenin results in its dissociation from E-cadherin (Roura *et al.*, 1999), verifies that tyrosine phosphorylation is an important mechanism for regulating the E-cadherin-catenin linkage. Tyrosine phosphorylation has also been reported to disrupt the β-catenin–α-catenin linkage (Ozawa and Kemler, 1998), although the critical tyrosine(s) in this case has not been determined. These observations suggest that multiple targets for tyrosine phosphorylation exist to regulate cell-cell adhesion.

The PTP Pez (PTPD2/PTP36) is a 130-kDa cytosolic (nontransmembrane) PTP (Smith *et al.*, 1995) expressed in a number of tissues. It is a member of the FERM (four-point-one,

ezrin, radixin, moesin) family of PTPs characterized by a conserved N-terminal FERM domain (Chishti *et al.*, 1998) and a C-terminal PTP catalytic domain separated by an intervening region. We recently showed that the subcellular localization of Pez is regulated in both HeLa and human umbilical vein endothelial cells (HUVEC); in cells grown to confluence Pez is localized to the cytosol, where it is concentrated at intercellular junctions, but it is predominantly nuclear in sparsely plated cells that have not yet formed extensive cell-cell contacts (Wadham *et al.*, 2000). Other factors also regulate the subcellular localization of Pez, including TGFβ, which inhibits translocation of Pez from the cytosol to the nucleus, and serum, which promotes the accumulation of Pez in the nucleus (Wadham *et al.*, 2000). Together these findings suggest that Pez could have multiple roles, involving the dephosphorylation of different substrates depending on whether it is in the nucleus or at intercellular junctions. Its presence at the intercellular junctions of confluent monolayers suggests that it may regulate the assembly or disassembly of adhesion complexes.

To elucidate the function of Pez, we used a "substrate trapping" approach (Flint *et al.*, 1997) in combination with the generation and overexpression of a dominant negative form of Pez to identify its substrates. We identified β-catenin as a substrate and show that the dominant negative Pez enhances both tyrosine phosphorylation of adherens junctions and cell motility.

MATERIALS AND METHODS

Tissue Culture and Cell Lines

MDCK and HEK 293 cell lines were cultured in DMEM supplemented with 10% FBS. HUVECS were obtained from discarded umbilical cords and cultured in M199 medium supplemented with 20% FBS and endothelial growth factors as previously described (Wall *et al.*, 1978). Transient transfections of HEK293 cells were carried out using Lipofectamine 2000 (Invitrogen Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. Stable MDCK clones were generated by transfection by standard calcium phosphate coprecipitation and transfectants selected by G418 (Promega, Madison, WI) resistance.

Antibodies

The polyclonal Pez antibody had previously been characterized (Wadham *et al.*, 2000). All other antibodies used were purchased: monoclonal Flag epitope antibody (M2) from AMRAD Biotech (Victoria, Australia), β-catenin, γ-catenin, and E-cadherin monoclonal antibodies from Transduction Laboratories, α-catenin mAb from either Transduction Laboratories (Lexington, KY) or Zymed (San Francisco, CA), monoclonal antiphosphotyrosine antibody (PY100) from New England Biolabs (Beverley, MA), and polyclonal ZO-1 and monoclonal p120catenin antibody from Zymed.

Generation of Mutant Pez Constructs

Isolation of the human Pez cDNA and generation of a Flag epitope-tagged construct in the mammalian expression vector, pcDNA3 (Invitrogen) has been described (Wadham *et al.*, 2000). The $D_{1079}A$ and R1127M mutations in Pez cDNA were made by site-directed mutagenesis using PCR. ΔFERM (amino acids 337-1187)- and ΔPTP (amino acids 1–932)-Pez were generated by PCR using the appropriate primers to remove the entire FERM or PTP domain, respectively. All constructs were tagged with the Flag epitope. The sequences of all mutated constructs were verified by sequencing from both the sense and antisense directions.

Generation of GST-Pez Fusion Proteins

wt-Pez and ST-Pez coding sequences were excised from the pcDNA3 constructs described above and cloned into the pGEX 4T-1 vector (Amersham Biosciences, Piscataway, NJ) to generate GST-fusion Pez proteins. The constructs were transformed into BL21-Codon Plus (DE3)-RIL *Escherichia coli* (Stratagene, La Jolla, CA) for protein expression. Cultures were induced with 0.15 mM IPTG for 2 h at ambient temperature, and the GST-fusion proteins were purified on glutathione sepharose. The amounts of full-length fusion proteins produced were determined by Coomassie blue staining after PAGE. Equal amounts of GST-wt-Pez and GST-ST-Pez protein were used for in vitro dephosphorylation of β-catenin.

Substrate Trapping

Newly confluent HUVEC lysates, used as a source of tyrosine-phosphorylated proteins, were prepared as described (Flint *et al.*, 1997). Briefly, the cells were incubated for 30 min with 50 μM sodium pervanadate to enrich for tyrosine-phosphorylated proteins, washed in phosphate-buffered saline (PBS), and lysed in ST buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 150 mM NaF, 10 mM sodium pyrophosphate, 5 mM EDTA, 1% Triton X-100, and protease inhibitor cocktail [P2714, Sigma, St Louis, MO]) containing 1 mM sodium orthovanadate at 4°C. The lysates were incubated on ice for 30 min in the presence of 5 mM iodoacetic acid to irreversibly inactivate endogenous PTPs. Unreacted iodoacetic acid was inactivated with 10 mM DTT. The lysates were then frozen on liquid nitrogen and stored at −70°C.

Flag-tagged wt-Pez or ST-Pez was transiently transfected into HEK293 cells and transfectants lysed in ST buffer 48 h after transfection. Equal amounts of protein from each lysate were immunoprecipitated (in the absence of orthovanadate) with an anti-Flag (M2) antibody precoupled to protein A sepharose beads. The Pez immunoprecipitates were washed three times in ST buffer, added to the phosphotyrosine enriched HUVEC lysates, and rocked at 4°C for 2 h. The beads were washed three times with ST buffer, boiled in Laemmli sample buffer, and bound proteins resolved by 8% SDS-PAGE. To detect tyrosine-phosphorylated proteins "pulled-down" by either wt or ST Pez, Western blots were performed using an antiphosphotyrosine antibody (PY100).

Immunoprecipitations and Western Blots

Immunoprecipitations were carried out after lysis of cells in ice-cold ST buffer containing 1 mM orthovanadate. Lysates were precleared with 20 μl of protein A-sepharose for 30 min at 4°C. Protein concentration was assayed using Bradford Reagent from Bio-Rad (Hercules, CA). Equal amounts (1–5 mg) of protein were incubated with 2 μg of primary antibodies supplemented with 20 μl packed protein A-sepharose for 1 h at 4°C. After washing, bound proteins were eluted by boiling in Laemmli sample buffer for 5 min separated by 8% SDS-PAGE and transferred to PVDF membrane (Hybond-P, Amersham Pharmacia Biotech) for Western blotting. Western blotting was carried out after blocking with 5% milk, 0.1% Triton X-100 in PBS using the indicated antibodies and developed using HRP-conjugated secondary antibody (Immunotech, Marseille, France) and ECL (Amersham Pharmacia Biotech). For quantitation, Western blots were developed with ECL-Plus (Amersham Pharmacia Biotech), and fluorescence intensity was imaged using a Molecular Dynamics Typhoon 9410 (Amersham Biosciences, United Kingdom) variable mode imager.

Immunofluorescence

MDCK stable cells lines expressing either wt-Pez, ΔPTP-Pez, or ΔFERM-Pez were plated at confluent density onto fibronectin

coated glass LabTek chamber slides (Nalge, Nunc International, Naperville, IL) and incubated for 2–3 d before staining. The cells were fixed in 4% paraformaldehyde/PBS for 10 min, quenched with 10 mg/ml sodium borohydride for 15 min, and then permeabilized by treatment with 0.1% Triton X-100. Primary antibodies were used at 1:100 dilution and binding detected by incubation with either fluorophore-coupled secondary antibodies or biotinylated secondary antibodies followed by fluorophore-conjugated streptavidin, as indicated (Molecular Probes, Eugene, OR).

Epifluorescence microscopy was performed on an Olympus BX-51 microscope equipped with excitation filters for Alexa Fluor 594/Texas red and fluorescein (494 nm), acquired to a Cool Snap FX, charge-coupled device (CCD) camera (Photometrics, Phoenix, AZ). Images were adjusted for brightness and contrast with V^{++} software (Digital Optics Ltd., Auckland, New Zealand). The line-profiling feature of this software was used to plot the intensity vs. position of different fluorophores along a path through the cell monolayer, in cells that had been costained for two proteins. Confocal microscopy was performed using a 60× oil-immersion objective on an Olympus IX70 inverted microscope linked to a Bio-Rad Radiance 2100 confocal microscope. Sequential scans of each fluorophore separately were carried out for two-color colocalization studies.

Wounding Assay

MDCK stable cell lines were plated onto six-well trays at densities that would give confluent monolayers after 24 h. Confluent monolayers were incubated a further 48 h to allow intercellular junctions to mature before being serum-starved for 24 h. A linear wound was generated on the monolayers by scraping with the edge of a cell scraper. Unattached cells were washed off with agitation. Cells were photographed at the same point on a grid at the time of scraping and again at 24 h later. The difference in width of the wound between the two edges at the time of scraping and 24 h later was measured and represents the distance migrated. Each line was plated and wounded in triplicate.

RESULTS

Pez Colocalizes with E-cadherin at AJs

We previously observed in confluent endothelial monolayers that endogenous Pez localizes to the intercellular junctions (Wadham *et al.*, 2000). Here, we investigate the localization of Flag-tagged-Pez stably expressed in MDCK cells, a polarized epithelial cell line in which cell-cell adhesion is well characterized. For subsequent investigations of the function of Pez, truncation mutants of Pez that lack either the catalytic (ΔPTP-Pez) or FERM (ΔFERM-Pez) domain (Figure 1A) were created and their subcellular localization when stably expressed in MDCK epithelial cell lines were also examined. wt-Pez and both truncated Pez mutants, examined by epifluorescent microscopy, were similarly localized to what appears to be the intercellular junctions (Figure 1B). To further confirm that the localization of Pez was indeed at intercellular junctions rather than at the cell surface, optical sectioning using a confocal microscope was performed on wt-Pez-MDCK cells that had been costained for the Flag-epitope (on Pez) and E-cadherin (a marker of AJs). The data showed that Pez precisely colocalized with E-cadherin at basolateral membranes both along the z-axis (Figure 1, C and D) and in the *x-y* plane (Figure 1E), confirming that it is localizing to the AJs.

ΔPTP-Pez Is a Potential Dominant Negative Mutant of Pez That Causes an Increase in Tyrosine Phosphorylation at AJs

Because ΔPTP-Pez is devoid of the catalytic domain and therefore not enzymatically active, but retains the ability to localize to intercellular junctions, it can potentially act as a dominant negative mutant. If Pez is an AJ PTP that regulates the level of tyrosine phosphorylation at AJs, then overexpression of a dominant negative mutant of Pez should result in an increase in tyrosine phosphorylation of AJ proteins. This was investigated using confluent monolayers of MDCK cells overexpressing ΔPTP-Pez. Cells were serum-starved followed by 10 min serum stimulation before staining with an antiphosphotyrosine antibody. Epifluorescence microscopy showed that there was a markedly higher level of tyrosine phosphorylation at intercellular junctions (marked by costaining with an anti–ZO-1 antibody) in ΔPTP-Pez–transfected cells compared with the empty vector control or wt-Pez transfectants (Figure 2A). Under these experimental conditions no tyrosine phosphorylation was detected at the intercellular junctions of empty vector- and wt-Pez–transfected cells. This is best demonstrated when the fluorescence intensities resulting from both the phosphotyrosine and ZO-1 antibodies were quantitated across several cell boundaries (Figure 2A, right column).

To determine the exact location of the tyrosine-phosphorylated proteins induced by overexpression of the putative dominant negative mutant ΔPTP-Pez, optical sectioning with a confocal microscope was performed on confluent ΔPTP-Pez–transfected cells costained with both the antiphosphotyrosine and anti–E-cadherin antibodies. As with Pez, the tyrosine phosphorylation induced by ΔPTP-Pez precisely colocalized with E-cadherin both along the Z-axis (Figure 2, B and C) and in the *x-y* plane, confirming that the increased tyrosine phosphorylation occurred at AJs. In addition, these data also suggest that the tyrosine-phosphorylated substrates remained in the proximity of the plasma membrane and did not translocate to other parts of the cell. Induction of tyrosine phosphorylation at intercellular junctions has also been confirmed using another potential dominant negative mutant, the R1127M point-mutant (Figure 2D; R1127 of Pez is the equivalent of R221 of PTP1B, which when mutated leads to inactivation of its PTP activity (Flint *et al.*, 1997).

The AJ Protein β-Catenin Is a Substrate of Pez and Coimmunoprecipitates with Pez

Data obtained so far suggested that Pez was an AJ PTP and in concordance with this hypothesis, a putative dominant negative mutant of Pez caused an increase in tyrosine phosphorylation of AJs. A number of components of the AJ complex can be tyrosine phosphorylated, leading to alterations in their functions (Steinberg and McNutt, 1999). We therefore used a substrate trapping approach to identify substrates of Pez at the AJ.

Asp181 of PTP1B is an essential residue for catalytic activity of PTP1B, which when mutated to Ala results in a catalytically inactive substrate trapping (ST) mutant (Flint *et al.*, 1997). Sequence alignment of the phosphatase domains of Pez and PTP1B indicate that Asp1079 of Pez corresponds to Asp181 of PTP1B. To generate a ST mutant of Pez (denot-

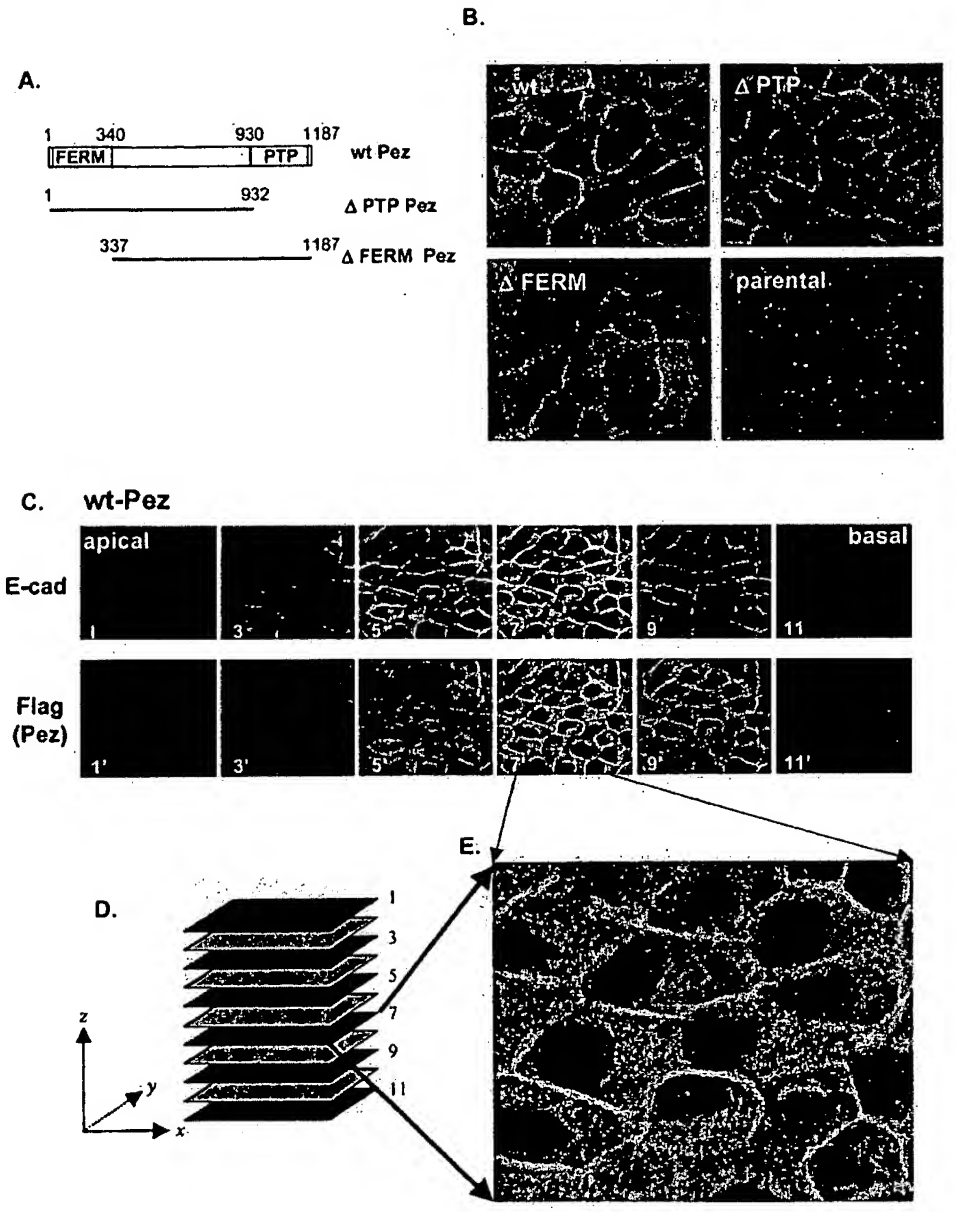


Figure 1. Flag-tagged wt-Pez and truncation mutants of Pez localize to the AJs of confluent monolayers of MDCK cells. (A) Schematic showing the structure of wt-, ΔPTP-, and ΔFERM-Pez. (B) Epifluorescence microscopy of confluent monolayers of parental MDCK cells and stable MDCK cell lines expressing wt-, ΔPTP-, and ΔFERM-Pez, labeled with an antibody against the Flag epitope and detected by indirect immunofluorescence using FITC-conjugated anti-mouse antibody. Specific labeling was observed in all Pez-expressing cell lines at the intercellular junctions. No staining was observed in the parental cells. (C) Z-series obtained by confocal laser scanning microscopy showing wt-Pez–transfected MDCK cells double-stained with the anti–E-cadherin (top row) and anti-Flag (bottom row) antibodies to indicate colocalization of Pez with E-cadherin along the z-axis. Anti–E-cadherin antibody was detected with phycoerythrin-conjugated anti-mouse IgG2a and anti-Flag antibody with biotinylated anti-mouse IgG1 antibody and FITC-conjugated streptavidin. Z-steps were carried out at 0.5-μm intervals, and alternate optical sections are shown. Top and bottom panels from each row show E-cadherin and Pez staining, respectively, from the same optical section. (D) Diagrammatic representation of the optical sections in the Z-series shown in C, indicating x, y, and z axes. (E) High-resolution merged image of one optical section (represented by panels 7 and 7′ from C) showing colocalization (yellow) in the x-y plane of Pez (green) and E-cadherin (red).

ed ST-Pez), we mutated D1079 of Pez to Ala and verified using an in vitro assay with a tyrosine-phosphorylated peptide that the catalytic activity of ST-Pez was significantly reduced compared with that of wt-Pez (between 10–20% of wt-Pez activity; unpublished data).

Without prior knowledge of the substrates of Pez, specific agonists could not be used to trigger their phosphorylation. Therefore, treatment with pervanadate, an inhibitor of PTPs, was used to upregulate tyrosine phosphorylation of proteins, including potential Pez substrates, in vivo (Figure 3A, panel 1). After lysis, endogenous PTPs in the HUVEC lysate were subsequently irreversibly inactivated by treatment with iodoacetic acid. The phosphotyrosyl-enriched HUVEC lysate devoid of endogenous PTP activities was incubated with either Flag-tagged wt- or ST-Pez (expressed in HEK 293 cells) immunoprecipitates bound to protein A-sepharose beads (Figure 3A, panel 6). Tyrosine-phosphorylated proteins that could interact with the wt- or ST-Pez immunoprecipitates, either directly or indirectly, were pulled-down and detected by Western blotting with an antiphosphotyrosine antibody. A number of tyrosine-phosphorylated proteins of similar staining intensities were pulled-down by both wt- and ST-Pez immunoprecipitates (Figure 3A, panel 2) but not immunoprecipitates from vector-transfected cells (panel 7). However, Band 1 was barely detectable in wt-Pez pull-downs but was clearly present when associated with ST-Pez. This could be because Band 1 did not

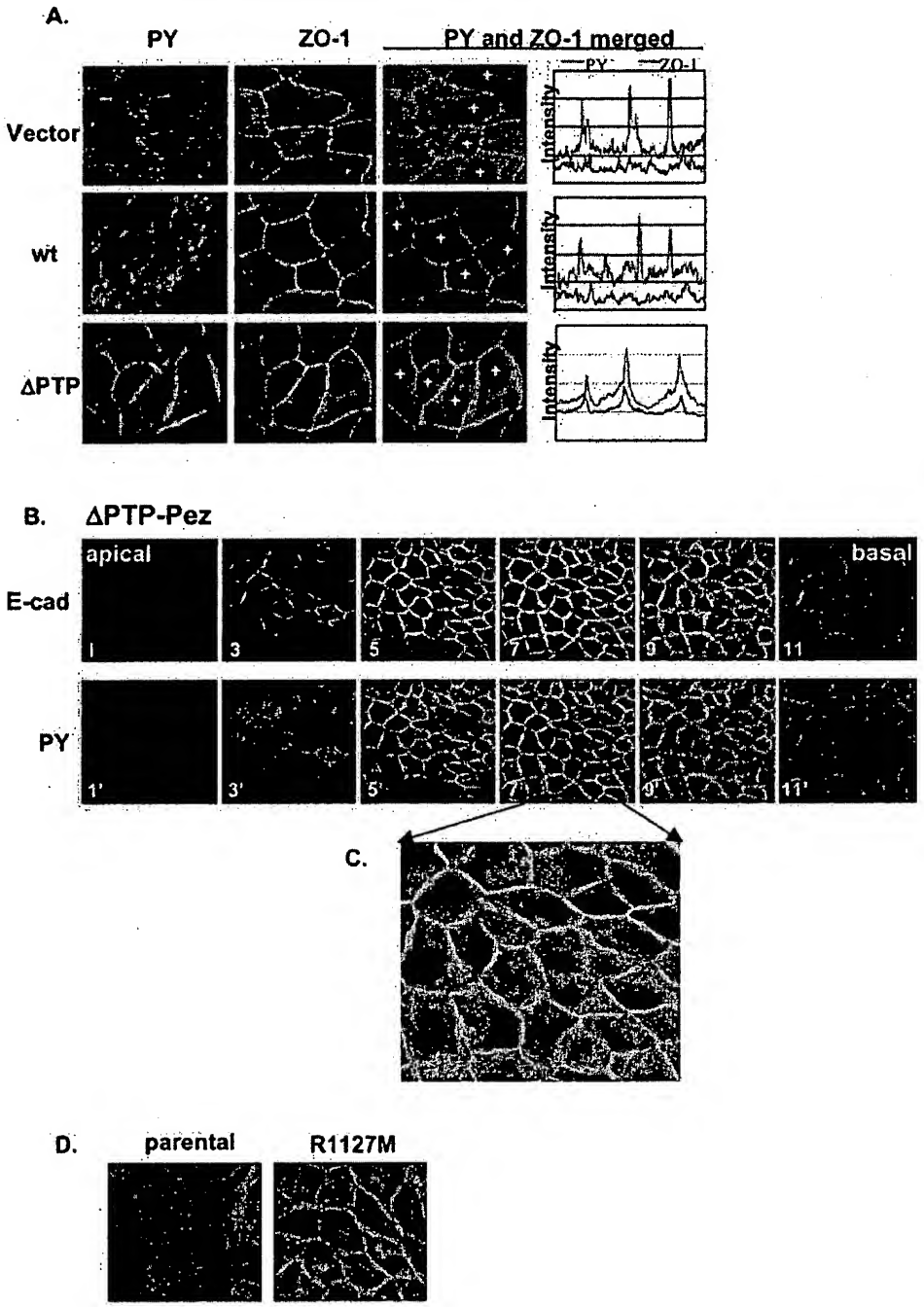


Figure 2. ΔPTP-Pez enhances tyrosine phosphorylation of AJs. (A) Epifluorescence microscopy showing confluent monolayers of MDCK cells stably expressing vector, wt- or ΔPTP-Pez, double-labeled with phosphotyrosine (PY) and ZO-1 antibodies and detected by indirect immunofluorescence using biotinylated anti-mouse antibody followed by Texas Red–conjugated streptavidin and Alexa fluor 350–conjugated anti-rabbit antibody, respectively. From left to right columns: phosphotyrosine antibody staining, ZO-1 staining to show positions of cell-cell contacts, merged image of PY staining (false colored red) and ZO-1 staining (false colored green), and quantitation of PY and ZO-1 fluorescent intensities taken along a line connecting the crosses shown in the merged image. (B) Z-series obtained by confocal laser scanning microscopy showing ΔPTP-Pez–transfected MDCK cells double-stained with the anti–E-cadherin (top row) and antiphosphotyrosine (bottom row) antibodies to show colocalization of phosphotyrosines with E-cadherin along the z-axis. Anti–E-cadherin antibody was detected with phycoerythrin-conjugated anti-mouse IgG2a, and antiphosphotyrosine antibody was detected with biotinylated anti-mouse IgG1 antibody and FITC-conjugated streptavidin. Z-steps were carried out at 0.5-μm intervals, and alternate optical sections are shown. Top and bottom panels from each row show E-cadherin and phosphotyrosine staining, respectively, from the same optical section. (C) High-resolution merged image of one optical section (represented by panels 7 and 7′ of B) showing colocalization (yellow) in the x-y plane of phosphotyrosine (green) and E-cadherin (red) staining. (D) Epifluorescence microscopy of confluent monolayers of parental MDCK or a stable MDCK cell-line expressing a Pez mutant, R1127M, stained with the phosphotyrosine antibody.

bind sufficiently stably to wt-Pez to be pulled-down or because the associated protein had been dephosphorylated by the catalytically active wt-Pez but not the inactive ST-Pez. Both causes for its absence in the wt-Pez pull-downs are consistent with Band 1 being a specific substrate of Pez.

Because the molecular weight of Band 1 is similar to that of the AJ protein β-catenin, we probed parallel lanes with a β-catenin antibody. β-Catenin was found to be pulled-down by both wt- and ST-Pez (Figure 3A, panel 3) and furthermore comigrated exactly with Band 1, suggesting that the substrate at this position could indeed be β-catenin. Thus, if indeed Band 1 is β-catenin, it was only tyrosine phosphorylated when in association with inactive ST-Pez, suggesting that the lack of tyrosine phosphorylation in the wt-Pez–associated protein was due to dephosphorylation by wt-Pez.

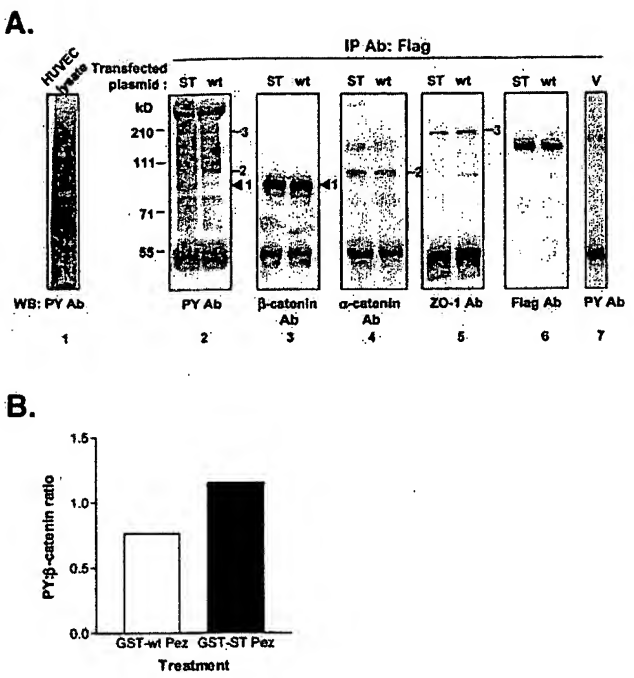


Figure 3. Identification of Pez substrates and interacting proteins by substrate trapping and by in vitro dephosphorylation. (A) Substrate trapping. HUVEC lysate enriched for tyrosine-phosphorylated proteins (panel 1) was incubated for 2 h with Sepharose-bound anti-Flag immunoprecipitates from cells transfected with wt-Pez, ST-Pez, or empty vector as shown. The beads were then washed, and bound proteins were eluted and immunoblotted with various antibodies as indicated. (B) In vitro dephosphorylation of β-catenin by GST-Pez. HUVEC lysate enriched for tyrosine-phosphorylated proteins (as in A) was incubated with either GST-wt-Pez or GST-ST-Pez bound to glutathione sepharose beads. After removal of the GST-Pez fusion proteins, β-catenin was immunoprecipitated, and the immunoprecipitates were Western blotted with an antiphosphotyrosine antibody. The amount of tyrosine-phosphorylated β-catenin was quantitated by fluorimaging. The blots were then stripped, counterblotted with anti–β-catenin antibody, and quantitated for total β-catenin. The ratio of tyrosine-phosphorylated β-catenin to total β-catenin for each GST-Pez treatment is shown (one representative of three experiments).

After stripping and reprobing the filters with antibodies to other components of junctional adhesion complexes, we identified a number of other junctional proteins, including α-catenin (band 2, Figure 3A, panel 4), ZO-1 (band 3, Figure 3A, panel 5), and plakoglobin (unpublished data), which interact with Pez but do not appear to be substrates for its PTPase activity. β-Catenin, α-catenin, and plakoglobin are all components of the AJ complex. ZO-1 is normally associated with tight junctions in polarized epithelial and endothelial monolayers but in newly confluent HUVEC, which have not formed bona fide tight junctions, it is also associated with AJs (Stevenson and Keon, 1998), suggesting that a complex of AJ proteins may be binding to Pez.

The phosphotyrosyl-enriched HUVEC lysate that was the source of phospho-β-catenin used in the substrate-trapping approach was made devoid of any active PTPs by iodoacetate treatment. However, there is a possibility that the wt-Pez immunoprecipitates incubated with the lysate may contain other active PTPs in addition to Pez. To demonstrate that phospho-β-catenin can be directly dephosphorylated by Pez, we carried out a similar experiment using recombinant Pez, prepared as a GST fusion protein from *E. coli*. As a control, we performed the same reaction with the inactive GST-ST-Pez. After incubation of recombinant Pez with the HUVEC lysate, β-catenin was immunoprecipitated with anti–β-catenin antibody and the amount of tyrosine-phosphorylated β-catenin remaining was quantitated by fluorimager analysis of phosphotyrosine immunoblots. Although the GST fusion proteins expressed poorly in *E. coli* (>90% of the products were degraded, presumably because the large fusion protein [~160 kDa] was poorly folded), the data showed that GST-wt-Pez, but not GST-ST-Pez, dephosphorylated the β-catenin in the lysate, removing about one third of the tyrosylphosphates on the β-catenin (Figure 3B). GST-Pez proteins lacking the FERM domain were also expressed and assayed, showing that the forms with wt but not ST catalytic domains dephosphorylate β-catenin to a similar extent as full-length GST-wt-Pez (unpublished data).

These experiments indicated that Pez can directly dephosphorylate β-catenin in vitro. To further assess whether Pez and β-catenin interact in vivo, we investigated whether Pez can be coimmunopreciptated with β-catenin. Immunoprecipitation of endogenous β-catenin from cell lysates containing either endogenous or ectopically expressed Pez was carried out, followed by Western blotting to determine if Pez is coimmunoprecipitated. Endogenous Pez was detected in β-catenin immunoprecipitates when a β-catenin antibody, but not a control antibody, was used to immunoprecipitate β-catenin from confluent HUVEC monolayers (Figure 4A). Similarly, ectopically expressed Pez coimmunoprecipitated with β-catenin in cells that were transfected with an expression vector bearing Pez cDNA but not empty vector (Fig, 4B), confirming that Pez and β-catenin interact in vivo.

Although β-catenin coimmunoprecipitated with both endogenous and ectopic Pez, the coimmunoprecipitation appeared to be relatively weak. This could have a number of possible explanations. First, the coimmunoprecipitations were carried out in 1% Triton X-100 with a brief pulse of sonication to maximize recovery of Pez. These conditions may be too harsh for the proteins to remain bound. In preliminary experiments to test the strength of the interaction between Pez and β-catenin, we have observed better coimmunoprecipitations if the sonication step was omitted although more striking increases in coimmunoprecipitation were observed by reducing the concentration of detergent used in the lysis buffer from 1% to 0.5% Triton X-100 (Figure 4B, right panel). This suggests that the complex formed may be detergent labile. Such detergent lability has been well documented with p120 catenin (p120ctn)-cadherin interactions, whereby only 5–20% of the p120ctn in detergent lysates is associated with cadherin in contrast to its almost complete localization to AJs or membrane fraction under detergent-free conditions (reviewed in Anastasiadis and Reynolds, 2000). Finally, it is likely that not all the Pez within the cell is localized to the AJ. This has certainly been observed in endothelial cells where Pez expression is observed in the cytosol away from the cell junctions, even when the monolayer is confluent (Wadham *et al.*, 2000). This is also

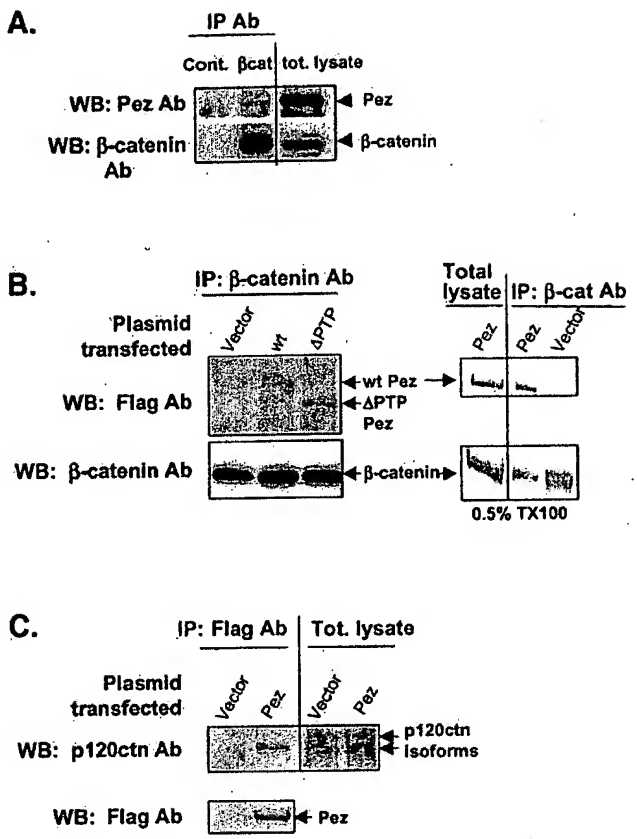


Figure 4. Coimmunoprecipitations of endogenous Pez and ectopically expressed wt- and ΔPTP-Pez with endogenous β-catenin and p120ctn. (A) HUVEC lysate was incubated with β-catenin antibody or control isotype-matched irrelevant antibody, and immunoprecipitates were Western blotted with Pez or β-catenin antibodies as indicated; right panel shows Western blots of total lysate. (B) Left panel: Empty vector, wt-Pez, or ΔPTP-Pez were transiently transfected into HEK293 cells. β-Catenin was immunoprecipitated from the lysates of transfectants and Western blotted using a Flag-epitope antibody to detect ectopically expressed Pez. The blot was stripped and reblotted with a β-catenin antibody showing the presence of β-catenin in all three immunoprecipitations. Right panel as in left panel, except that lysis was carried out in 0.5% instead of the 1% Triton X-100, which was used in all the other immunoprecipitations shown. (C) Flag epitope-tagged Pez was immunoprecipitated with an anti-Flag antibody from HEK 293 cells stably transfected with either empty vector or Flag epitope-tagged Pez. The immunoprecipitates were Western-blotted with an anti-p120ctn antibody and reblotted with the anti-Flag antibody.

particularly true when Pez is ectopically expressed in HEK 293 cells (unpublished data).

Because a number of AJ proteins were also pulled-down together with β-catenin (Figure 3A) and because of their structural relatedness, we also investigated whether p120ctn coimmunoprecipitates with Pez. Immunoprecipitates from lysates of vector- or Pez-transfected stable HEK 293 cell lines carried out using the Flag-epitope antibody were Western blotted with a p120ctn antibody. p120ctn coimmunoprecipitated with Pez that was immunoprecipitated with the Flag antibody (Figure 4C). Western blots of total lysates from both the vector- and Pez-transfected cells showed that HEK 293 cells expressed two major p120ctn isoforms of similar abundance (Figure 4C), with the larger isoform corresponding to full-length p120ctn (isoform 1). Interestingly, only the smaller MW isoform (~95 kDa) coimmunoprecipitated with Pez. Isoforms 1 and 3 are the most commonly expressed (reviewed in Anastasiadis and Reynolds, 2000), hence the smaller isoform coimmunoprecipitating with Pez is most likely isoform 3, although this remains to be confirmed. It is unclear why Pez is only associated with one isoform of p120ctn, and it is also unkown at this stage whether p120ctn interacts directly with Pez and whether it is a substrate. These will be the subject of future studies. What is clear, however, is that the cell junctional proteins that are pulled-down with Pez are likely to be specific because the coimmunoprecipitation discriminated between the two highly related isoforms of p120ctn.

ΔPTP-Pez Interacts with and Induces Tyrosine Phosphorylation of β-Catenin

If ΔPTP-Pez acts as a dominant negative mutant of Pez (as suggested by its localization to the AJ [Figure 1B] and by its ability to induce tyrosine phosphorylation of proteins at the AJ [Figure 2]), and if β-catenin is a bona fide substrate of Pez, then ΔPTP-Pez should interact with β-catenin in vivo to increase its tyrosine phosphorylation. To investigate the ability of ΔPTP-Pez to interact with β-catenin in vivo, β-catenin was immunoprecipitated from HEK293 cells stably transfected with ΔPTP-Pez. The β-catenin immunoprecipitates were then Western blotted with an anti-Flag antibody to detect ΔPTP-Pez that has coimmunoprecipitated with β-catenin. The data showed that ΔPTP-Pez was able to coimmunoprecipitate with β-catenin to the same extent as wt-Pez (Figure 4B, left panel).

To examine whether ΔPTP-Pez can induce tyrosine phophorylation of Pez substrates through a dominant negative effect, antiphosphotyrosine Western blots were performed on extracts from confluent monolayers of MDCK cells stably expressing empty vector, wt-Pez, or ΔPTP-Pez. Cells were serum-starved for 24 h before addition of serum for 10 min to induce tyrosine phosphorylation. To see specific tyrosine phosphorylation of Pez substrates, cells were not pretreated with pervanadate, which would have caused a global increase in tyrosine phosphorylation. Thus, in concordance with other studies (Ayalon and Geiger, 1997), the basal level of tyrosine-phosphorylated proteins in vector-transfected cells was very low. Two bands that were specifically phosphorylated in extracts from ΔPTP-Pez but not empty vector or wt-Pez transfected cells (Figure 5A, closed arrowheads) were observed. The lower MW band comigrated with β-catenin when the filter was counterblotted with a β-catenin antibody, suggesting that one of the proteins that is tyrosine phosphorylated through overexpression of ΔPTP-Pez is β-catenin. Immunoprecipitation of β-catenin followed by Western blotting with an antiphosphotyrosine antibody confirmed that β-catenin was indeed phosphorylated in ΔPTP-Pez–transfected but not empty vector-transfected cells (Figure 5B). These data indicate that ΔPTP-Pez acts as a dominant negative mutant of Pez leading to an increased level of β-catenin tyrosine phosphorylation. The finding that ΔPTP-Pez could interact with and increase

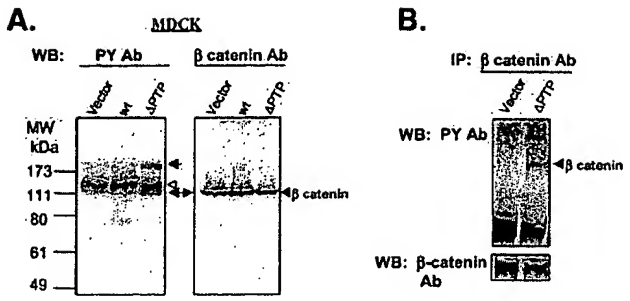


Figure 5. ΔPTP-Pez enhances tyrosine phosphorylation of β-catenin. (A) MDCK cell lines stably expressing vector, wt-Pez, or ΔPTP-Pez were serum-starved for 24 h followed by 10-min stimulation with 10% FBS to induce tyrosine phosphorylation. Lysates were Western blotted with a phosphotyrosine (PY) antibody (left). Arrowheads show two proteins specifically tyrosine phosphorylated in the ΔPTP-Pez transfected cells, one of which comigrates with β-catenin (double arrowhead) when the blot was stripped and reblotted with a β-catenin antibody (right). (B) Lysates from empty vector- or ΔPTP-Pez–transfected MDCK cells were immunoprecipitated with a β-catenin antibody followed by Western blotting of the immunoprecipitates with an antiphosphotyrosine antibody (top). The blot was stripped and counterprobed with a β-catenin antibody (bottom).

the tyrosine phosphorylation status of β-catenin further reinforces our conclusion that β-catenin is a bona fide Pez substrate. The presence of another protein with increased tyrosine phosphorylation in ΔPTP-Pez but not wt-Pez or empty vector-transfected cells (Figure 5A, open arrowhead) suggests that there are other Pez substrates that were not identified by substrate trapping. Any additional substrates are likely to also be located at the intercellular junctions because the tyrosine phosphorylation induced by ΔPTP-Pez was highly specific to intercellular junctions (Figure 2A). The presence of other substrates at intercellular junctions in addition to β-catenin would also account for the dramatic increase in tyrosine phosphorylation at intercellular junctions in ΔPTP-Pez MDCK cells. Similarly, the tyrosine-phosphorylated band comigrating with β-catenin in Figure 5A may be comprised of more than one protein of the same relative mobility. The higher degree of tyrosine phosphorylation relative to β-catenin protein in the ΔPTP-Pez lysate shown in Figure 5A compared with that in the β-catenin immunoprecipitates in Figure 5B (where the observed phosphorylation is due solely to β-catenin) suggests that this may be the case.

Overexpression of the Dominant Negative Mutant (ΔPTP-Pez) Enhances Cell Migration

Tyrosine phosphorylation of β-catenin has been correlated with increased cell migration in a number of studies (Liu *et al.*, 1997;Muller *et al.*, 1999;Hollande *et al.*, 2001). We used an in vitro "wound" assay to investigate whether the increase in tyrosine phosphorylation of β-catenin that results from overexpression of the dominant negative mutant, ΔPTP-Pez, could affect rates of cell migration. In this assay, a linear scratch was made on a confluent monolayer of MDCK cells to generate a linear denuded area, after which cells from the edge of the wound could migrate into the denuded area to repopulate it. After 24 h, cells overexpressing ΔPTP-Pez had migrated further into the wound than cells overexpressing either empty vector or wt-Pez (Figure 6A). Measurements of the distances migrated after 24 h (Figure 6B) showed that the average distance migrated by the ΔPTP transfectants was significantly greater ($p = 0.02$) than the distance migrated by either wt-Pez or vector control cells. There was no significant difference between the distances migrated by the wt-Pez cells and empty vector-transfected cells ($p = 0.08$). Higher resolution images of the "wound" shows that the cells at the edge are migrating into the wound characterized by protrusions into the wound and formation of pseudopodia (Figure 6C). Our observation that the dominant negative mutant, ΔPTP-Pez, enhanced cell motility suggests that Pez is a regulator of cell motility, most likely through its role in regulating cell-cell adhesion.

DISCUSSION

Regulation of the integrity of AJs in response to external cues is important for proper tissue and organ formation during embryonal development and for wound healing in adults. The participation of tyrosine phosphorylation in regulation of AJ function is apparent from the involvement of PTKs such as Src (Behrens *et al.*, 1993;Fujita *et al.*, 2002), EGF receptor, and HGF/scatter factor receptor (Shibamoto *et al.*, 1994) in phosphorylation of AJ proteins, leading to decreased cell-cell interaction and concomitant enhancement of cell migration. A number of PTPs have also been found to be components of the AJ complex, including the receptor PTPs μ, κ, and LAR and the cytosolic PTP, PTP1B (reviewed in (Steinberg and McNutt, 1999).

In this study, we have identified the PTP Pez as a novel PTP of AJs. A truncation mutant of Pez lacking the catalytic domain acted as a dominant negative mutant to enhance tyrosine phosphorylation of AJs and promote cell migration. Analysis of the proteins that are tyrosine phosphorylated as a result of the overexpression of the dominant negative mutant suggested that there are at least two Pez substrates in epithelial cells, one of which is the AJ protein β-catenin. Using a substrate trapping mutant to isolate potential Pez substrates followed by in vitro dephosphorylation of β-catenin by recombinant Pez, we confirmed that β-catenin is indeed a substrate of Pez. Dephosphorylation of β-catenin by recombinant Pez in the absence of any other active PTP also demonstrated that Pez could directly dephosphorylate β-catenin. Both endogenous and ectopically expressed Pez coimmunoprecipitated with endogenous β-catenin, indicating that they interact in vivo, providing further evidence that Pez is a physiological regulator of β-catenin tyrosine phosphorylation.

The highly similar complement of AJ proteins pulled-down by both wt-Pez and ST-Pez, the observation that wt-Pez and ST-Pez can pull-down β-catenin equally well, and the observation that wt-Pez can pull-down unphosphorylated β-catenin all suggest that Pez is likely to be a component of the AJ complex. This raises a number of questions. First, although we have demonstrated that Pez could directly dephosphorylate β-catenin, the question remains as to whether its association with the AJ complex is solely through binding to β-catenin or through binding to some

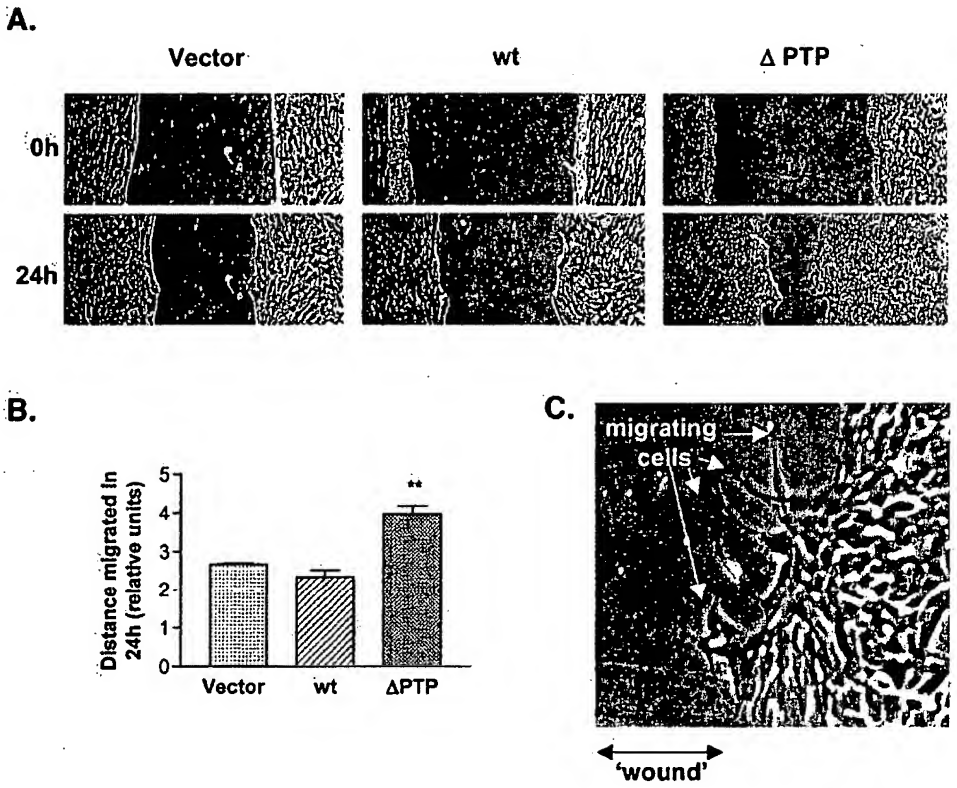


Figure 6. ΔPTP-Pez enhances cell motility. (A) Confluent monolayers of MDCK cells stably expressing empty vector, wt-Pez, or ΔPTP-Pez were wounded using the edge of a cell scraper and photographed immediately (0h). Each wound was photographed again at the same spot 24 h later. (B) The distance between the wound edges was measured at the same point on each wound at 0 h and at 24 h. The difference in distance between the two edges at 24 h and 0 h was taken to be the distance migrated in 24 h. Each cell line was plated and wounded in triplicate and the experiment performed at least twice with similar outcomes; data from one representative experiment are presented. **p = 0.02 when distance migrated by ΔPTP- expressing cells was compared with cells expressing empty vector. (C) Higher magnification image of ΔPTP-Pez MDCK cells at the wound edge 24 h after wounding.

other component of the AJ complex. Second, if Pez is a component of the AJ through direct binding to β-catenin or some other protein, than one might expect that any tyrosine phosphorylation of β-catenin at the AJ will be very transient. To achieve longer-term loosening of the AJ might then require downregulation of Pez activity.

Of the PTPs that have previously been shown to be localized to the AJ, PTP LAR and PTP 1B have been shown to dephosphorylate β-catenin, whereas the substrates of PTPμ are yet to be identified. There are potentially many reasons for multiple PTPs to be associated with AJs. These include cell type-specific expression of some PTPs, different degrees of responsiveness to external stimuli and different substrate specificities exhibited by different PTPs. In the case of β-catenin, the crystal structure indicates there are potentially up to 14 tyrosines that are accessible for phosphorylation. One of these, Tyr654, has been demonstrated to regulate the binding of β-catenin to E-cadherin (Roura *et al.*, 1999), but the phospho-tyrosine that interrupts α-catenin binding is yet to be determined. It is conceivable that dephosphorylation of different tyrosines that mediate different functions on the one molecule may be regulated by different PTPs. A comprehensive analysis of substrate specificity and responses to external stimuli for individual PTPs, which to date has not been carried out, is essential to fully elucidate the role of tyrosine phosphorylation in regulating AJ functions.

Finally, some studies have reported that after the tyrosine phosphorylation of β-catenin and its dissociation from intercellular junctions, it is translocated into the nucleus where it can, under some circumstances, interact with the LEF-1/Tcf transcription factor to alter gene expression (Adam *et al.*, 2001; Kim and Lee, 2001; Monga *et al.*, 2002). What is not clear from these studies is whether β-catenin is translocated into the nucleus in its tyrosine-phosphorylated form and if so, whether the tyrosine-phosphorylated form can interact with LEF-1/Tcf. Intriguingly, under conditions where cell-cell adhesion is disrupted, we have shown that Pez also translocates into the nucleus (Wadham *et al.*, 2000). It would be particularly important to determine whether Pez and β-catenin interact in the nucleus and if so, what the functional consequence of this interaction is.

ACKNOWLEDGMENTS

We thank Jenny Drew and Anna Sapa for the preparation and culture of HUVEC, Dr. Ghafar Sarvestani for assistance with confocal microscopy, the staff of the Women's and Children's Hospital and Burnside Hospital, Adelaide, South Australia for the collection of umbilical cords, and Drs. B.W. Wattenberg and G.J. Goodall for reading the manuscript. This work was supported in part by grants from the Royal Adelaide Hospital to Y.K.-G. and a National Health and Medical Research Council Program Grant to M.A.V and J.R.G. C.W is the recipient of a Dawes Postgraduate Research Scholarship from the Royal Adelaide Hospital, Australia.

REFERENCES

Adam, L., Vadlamudi, R.K., McCrea, P., and Kumar, R. (2001). Tiam1 overexpression potentiates heregulin-induced lymphoid enhancer factor-1/beta-catenin nuclear signaling in breast cancer cells

by modulating the intercellular stability. J. Biol. Chem. *276*, 28443–28450.

Anastasiadis, P.Z., and Reynolds, A.B. (2000). The p120 catenin family: complex roles in adhesion, signaling and cancer. J. Cell Sci. *113* (Pt 8), 1319–1334.

Ayalon, O., and Geiger, B. (1997). Cyclic changes in the organization of cell adhesions and the associated cytoskeleton, induced by stimulation of tyrosine phosphorylation in bovine aortic endothelial cells [In Process Citation]. J. Cell Sci. *110*, 547–556.

Behrens, J., Vakaet, L., Friis, R., Winterhager, E., Van Roy, F., Mareel, M.M., and Birchmeier, W. (1993). Loss of epithelial differentiation and gain of invasiveness correlates with tyrosine phosphorylation of the E-cadherin/beta-catenin complex in cells transformed with a temperature-sensitive v-SRC gene. J. Cell Biol. *120*, 757–766.

Chishti, A.H. *et al.* (1998). The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane. Trends Biochem. Sci. *23*, 281–282.

Cowin, P., and Burke, B. (1996). Cytoskeleton-membrane interactions. Curr. Opin. Cell Biol. *8*, 56–65.

Flint, A.J., Tiganis, T., Barford, D., and Tonks, N.K. (1997). Development of "substrate-trapping" mutants to identify physiological substrates of protein tyrosine phosphatases. Proc. Natl. Acad. Sci. USA *94*, 1680–1685.

Fujita, Y., Krause, G., Scheffner, M., Zechner, D., Leddy, H.E., Behrens, J., Sommer, T., and Birchmeier, W. (2002). Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. Nat. Cell Biol. *4*, 222–231.

Gumbiner, B.M. (1995). Signal transduction of beta-catenin. Curr. Opin. Cell Biol. *7*, 634–640.

Hollande, F., Choquet, A., Blanc, E.M., Lee, D.J., Bali, J.P., and Baldwin, G.S. (2001). Involvement of phosphatidylinositol 3-kinase and mitogen-activated protein kinases in glycine-extended gastrin-induced dissociation and migration of gastric epithelial cells. J. Biol. Chem. *276*, 40402–40410.

Kim, K., and Lee, K.Y. (2001). Tyrosine phosphorylation translocates beta-catenin from cell→cell interface to the cytoplasm, but does not significantly enhance the LEF-1-dependent transactivating function. Cell Biol. Int. *25*, 421–427.

Liu, D., el Hariry, I., Karayiannakis, A.J., Wilding, J., Chinery, R., Kmiot, W., McCrea, P.D., Gullick, W.J., and Pignatelli, M. (1997). Phosphorylation of beta-catenin and epidermal growth factor receptor by intestinal trefoil factor. Lab. Invest. *77*, 557–563.

Monga, S.P., Mars, W.M., Pediaditakis, P., Bell, A., Mule, K., Bowen, W.C., Wang, X., Zarnegar, R., and Michalopoulos, G.K. (2002). Hepatocyte growth factor induces Wnt-independent nuclear translocation of beta-catenin after Met-beta-catenin dissociation in hepatocytes. Cancer Res. *62*, 2064–2071.

Muller, T., Choidas, A., Reichmann, E., and Ullrich, A. (1999). Phosphorylation and free pool of beta-catenin are regulated by tyrosine kinases and tyrosine phosphatases during epithelial cell migration. J. Biol. Chem. *274*, 10173–10183.

Ozawa, M., and Kemler, R. (1998). Altered cell adhesion activity by pervanadate due to the dissociation of alpha-catenin from the E-cadherin. catenin complex. J. Biol. Chem. *273*, 6166–6170.

Roura, S., Miravet, S., Piedra, J., Garcia, d. H., Dunach, M. (1999). Regulation of E-cadherin/Catenin association by tyrosine phosphorylation. J. Biol. Chem. *274*, 36734–36740.

Shibamoto, S., Hayakawa, M., Takeuchi, K., Hori, T., Oku, N., Miyazawa, K., Kitamura, N., Takeichi, M., and Ito, F. (1994). Tyrosine phosphorylation of beta-catenin and plakoglobin enhanced by hepatocyte growth factor and epidermal growth factor in human carcinoma cells. Cell Adhes. Commun. *1*, 295–305.

Smith, A.L., Mitchell, P.J., Shipley, J., Gusterson, B.A., Rogers, M.V., and Crompton, M.R. (1995). Pez: a novel human cDNA encoding protein tyrosine phosphatase- and ezrin-like domains. Biochem. Biophys. Res. Commun. *209*, 959–965.

Steinberg, M.S., and McNutt, P.M. (1999). Cadherins and their connections: adhesion junctions have broader functions. Curr. Opin. Cell Biol. *11*, 554–560.

Stevenson, B.R., and Keon, B.H. (1998). The tight junction: morphology to molecules. Annu. Rev. Cell Dev. Biol. *14*, 89–109.

Vleminckx, K., and Kemler, R. (1999). Cadherins and tissue formation: integrating adhesion and signaling. Bioessays *21*, 211–220.

Wadham, C., Gamble, J.R., Vadas, M.A., and Khew-Goodall, Y. (2000). Translocation of protein tyrosine phosphatase Pez/PTPD2/PTP36 to the nucleus is associated with induction of cell proliferation. J. Cell Sci. *113* (Pt 17), 3117–3123.

Wall, R.T., Harker, L.A., Quadracci, L.J., and Striker, G.E. (1978). Factors influencing endothelial cell proliferation in vitro. J. Cell Physiol. *96*, 203–213.